



HAL
open science

Etudes structurales et fonctionnelles de l'IRES du VHC en association avec le motif de reconnaissance à l'ARN de la sous-unité b du facteur eIF3.

Julien Perard

► **To cite this version:**

Julien Perard. Etudes structurales et fonctionnelles de l'IRES du VHC en association avec le motif de reconnaissance à l'ARN de la sous-unité b du facteur eIF3.. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2009. Français. NNT: . tel-00436687

HAL Id: tel-00436687

<https://theses.hal.science/tel-00436687>

Submitted on 27 Nov 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

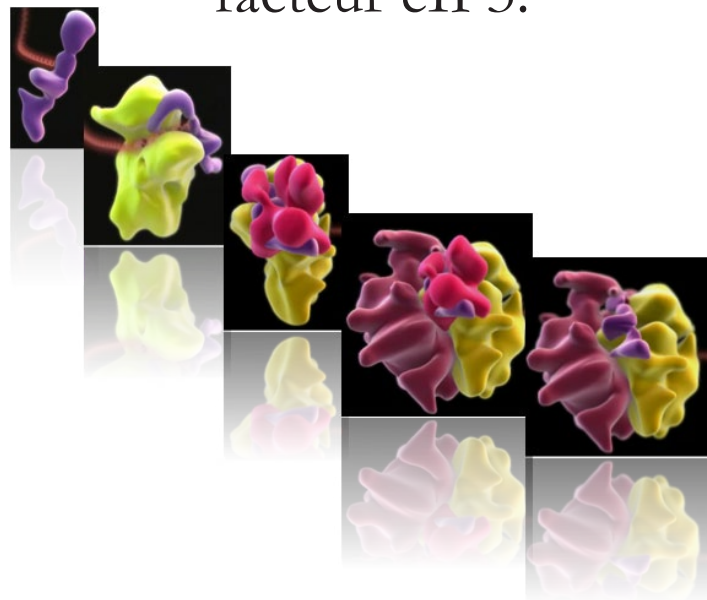
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE d'Université

Pour l'obtention du diplôme de Docteur de l'Université Joseph
Fourier-Grenoble I

Présenté par : **Julien PERARD**

Etudes structurales et fonctionnelles de l'IRES
du VHC en association avec le motif de
reconnaissance à l'ARN de la sous-unité b du
facteur eIF3.



Soutenue publiquement le 13 Novembre 2009 devant le jury :

Président du jury : Pr. Weissenhorn

Rapporteurs : Dr. Penin
Dr. Ohlmann

Examineur : Dr. Eriani

Directeurs de thèse : Dr. Baudin
Pr. Drouet

Travaux de thèse effectués à l'UVHCI (Unit of Virus Host Cell Interactions).
UMI 3265 CNRS-UJF-EMBL, Grenoble, France.

Ecole Doctorale (EDISCE)

INGÉNIERIE POUR LA SANTÉ , LA COGNITION ET L'ENVIRONNEMENT

PREFACE

«Chacun de nous, à vingt ans, a rêvé de transformer le monde. Chacun de nous sait à quarante ans qu'il ne le fera pas. Au mieux peut-il espérer, avec de la chance, apporter quelque contribution au capital de vérités, de recettes et d'idées que l'homme amasse lentement.»

François Jacob

RAPPEL DES LOIS DE MURPHY

Avant d'étudier quelque chose, il est recommandé de connaître entièrement son sujet. • Chaque solution amène de nouveaux problèmes • Dans le doute, soyez convaincant. • Dans toute série de mesures, celle qui paraît la plus juste est la plus fausse. • Des attentes négatives conduisent à des résultats négatifs. • Des attentes positives conduisent à des résultats négatifs • Dès que quelque chose fonctionne, toute amélioration l'empêche de fonctionner • Il est impossible de faire la différence entre la magie et une technologie de pointe • Il est parfaitement impossible de faire quoi que ce soit à l'épreuve des imbéciles, les imbéciles se montrent toujours si ingénieux ! • L'erreur est humaine - l'attribuer à quelqu'un d'autre que soi est encore plus humain • La nature nous réserve toujours des pannes • (i) Si c'est vert ou si ça remue, c'est de la biologie. (ii) Si ça pue, c'est de la chimie. (iii) Si ça ne fonctionne pas, c'est de la physique. (iv) Si ça occupe des tableaux entiers de formules, c'est des mathématiques • Laissées à elles mêmes, les choses tendent à aller de mal en pis • Le modèle que vous voulez n'est jamais celui qui est à vendre • Les expériences doivent être reproductibles : elles doivent toutes planter de la même façon • Les gens accepteront plus facilement votre idée si vous leur dites que Benjamin Franklin l'a eue avant vous • Méfiez-vous de l'homme qui travaille dur pour apprendre quelque chose, l'apprend, et ne se trouve pas plus intelligent qu'avant, il est plein d'un ressentiment amer vis à vis de la société, car elle contient beaucoup de gens qui n'ont pas eu à travailler dur pour rester ignorants • Ne jamais recommencer une expérience qui a fonctionné • Peu importe la quantité de travail réalisé, elle n'est jamais suffisante • Peu importe où est l'erreur, quelqu'un savait qu'il y en aurait une • Pour découvrir quelque chose, il faut être en train de chercher quelque chose. • Quel que soit le résultat prévu, il se trouve toujours quelqu'un pressé de : (i).mal l'interpréter; (ii).le critiquer; (iii).croire que cela provient directement de sa propre théorie • Quoi que vous décidiez de faire, il y a autre chose à faire auparavant • Résoudre un problème le change • Rien n'est aussi simple qu'il y paraît. • S'il existe plusieurs façons pour les choses de mal se passer, celle qui va arriver est aussi celle qui causera le plus de dégâts. Corollaire: C'est au pire moment que les choses tourneront mal. • Même si quelque chose ne peut pas aller mal, cela ira mal quand même • Si ça marche du premier coup, c'est que vous n'avez aucune idée de ce que vous avez fait • Si jamais vous décidez de faire quelque chose, il y aura toujours quelque chose d'autre à faire avant. • Si quelque chose ne peut tout simplement pas mal tourner, ca se tournera mal quand même. • Si tout semble manifestement bien se passer, c'est que quelque chose vous a échappé. • Si une expérience marche, c'est que quelque chose ne va pas. • Si vous avez décelé quatre façons possibles pour les événements de mal tourner, et si vous les avez circonvenues, alors une cinquième apparaîtra spontanément. • Si vous faites suffisamment de recherches, elles tendront à prouver votre théorie • Toujours conserver une trace de ses mesures - cela indique que vous avez travaillé. • Tout ce qui est susceptible de mal tourner tournera nécessairement mal. • Tout ce qui peut aller mal ira mal • Tout prend plus de temps que ce que vous pouvez prévoir. • Toutes les grandes découvertes sont faites par erreur • Tracez la courbe avant de placer les points • Avant même de fonctionner, tout programme est déjà obsolète. • Chaque programme non trivial est bogué. Corollaire : une condition suffisante pour qu'un programme soit trivial est qu'il ne contienne pas de bogues. • De nouveaux systèmes génèrent de nouveaux problèmes • L'erreur est humaine, mais pour empirer il faut un ordinateur • La complexité

de tout programme croît jusqu'à excéder les capacités du programmeur chargé de sa maintenance. • La plus grave erreur d'un programme ne peut pas être découverte avant que le produit soit en vente depuis six mois. • N'importe quel programme finira par occuper toute la mémoire. • Pour estimer combien de temps prendra l'écriture et le débogage d'un programme, prendre son estimation la plus large, multiplier par deux, rajouter un, et passer à l'unité de temps supérieure. • Quand rien ne marche, il est recommandé de lire les instructions. • Quand un compilateur accepte un programme sans la moindre erreur au premier essai, le programme ne donnera pas les résultats attendus. • Si l'on faisait en sorte que les programmeurs puissent programmer en Français, on se rendrait compte qu'ils ne savent pas écrire en Français. • Si le produit a été conçu pour rejeter toute entrée invalide, alors un idiot ingénieux découvrira une méthode pour rentrer quand même des données invalides. • Si le système de test fonctionne parfaitement, tous les autres systèmes auront des problèmes. • Si un programme est inutile, il faudra augmenter la documentation le concernant. • Si un programme est utile, il faudra le changer. • Tout plantera au moins une fois. • Un programme informatique fait ce que vous lui avez dit de faire, pas ce que vous voulez qu'il fasse.

REMERCIEMENTS

Je remercie le directeur de l'UMI-UVHCI, le Dr. Stephen Cusack, ainsi que le Pr. Rob Ruigrok pour m'avoir accueilli dans cet institut durant mes trois années de thèse où j'ai pu trouver un environnement scientifique d'une grande richesse aussi bien technique qu'intellectuelle me permettant de mener à bien mes travaux.

Je tiens à remercier mes deux directeurs de thèse, le Pr. Emmanuel Drouet et le Dr. Florence Baudin. Emmanuel, merci pour tes conseils et ton encouragement. Florence merci, pour tes FBA et ta patience, sans toi cette thèse aurait eu un autre visage.

Tous mes remerciements vont également au Dr. Jérôme Boisbouvier pour avoir embelli ce sujet et apporté ces quelques spins à l'IRES du VHC. Merci également à toute l'équipe du laboratoire de RMN et plus particulièrement à Isabel, Lionel et Maria pour leur patience et leur ténacité dans la production de ces échantillons si précieux.

Merci aux membres du jury d'avoir accepté de lire et de critiquer ce manuscrit.

Merci au Pr. Marc Jamin pour son aide dans le retraitement des données de SAXS et pour la modélisation de mes molécules.

Merci à tous les membres de l'UVHCI, les étudiants pour la bonne humeur dans les bureaux, les chercheurs pour leur patience et leurs intérêts, au personnel technique qui a toujours été présent, Pierre, désolé encore pour les litres de LB à autoclaver. Lucy, grâce à tes gâteaux et tes pauses café, j'ai pu me relaxer pendant les moments de stress, Alex, merci pour tes « Alexades » qui sont dignes d'entrer dans l'histoire. Merci également à Cédric et Ambroise pour leurs scripts magiques et pour l'aide sur les données de SAXS.

Merci à mes parents, qui ont dû me supporter pendant ces quelques années. Grâce à vous, j'ai pu aller au delà de mes rêves, et vous m'avez accompagné jusqu'au bout dans cette aventure, alors merci.

Je tiens à remercier tout particulièrement Isabel, ma femme, ma confidente et ma conseillère. Je te remercie du fond du cœur, tu as été patiente et constamment à l'écoute pour mes petits problèmes de «manips». Tu as toujours su trouver les mots pour m'encourager et me permettre d'aller de l'avant. Sans toi, cette thèse ne serait pas, alors merci.

A Sarah, ma fille, qui a su être si patiente, merci pour tes sourires et ta gentillesse qui m'ont suivi pendant cette dernière année de thèse.

TABLE DES MATIERES

PROLOGUE		1
INTRODUCTION		5
I	Les hépatites virales	7
II	Présentation détaillée de l'hépatite C	10
	II-1 Historique	10
	II-2 Epidémiologie	12
	II-3 Transmission	12
	II-4 Maladie et traitements	13
III	Le virus de l'hépatite C	15
	III-1 Classification	15
	III-2 Structure du VHC	16
	III-3 Organisation du génome et maturation de la polyprotéine	17
	III-4 Cycle de vie du VHC	19
	III-5 Modèles pour l'étude du VHC	19
IV	Initiation de la traduction dépendante de la coiffe	21
	IV-1 Expression génique et traduction chez les Eucaryotes	21
	IV-2 L'initiation et la reconnaissance de la coiffe de l'ARNm	21
	IV-2-A Formation du complexe ternaire eIF2•GTP•ARNtiMet	21
	IV-2-B Formation du complexe de pré-initiation 43S	21
	IV-2-C La coiffe est reconnue par le complexe eIF4F	23
	IV-2-D Le balayage	23
	IV-2-E Assemblage du ribosome 80S et fin de l'initiation	23
	IV-3 L'élongation	24
	IV-4 La terminaison	24
V	Initiation de la traduction CAP indépendante	24
VI	Un facteur commun : eIF3	27
VII	Les IRES	30
	VII-1 Les IRES cellulaires	31
	VII-2 Les IRES viraux	32
	VII-2-A Caractéristiques et diversités structurales	32
	VII-2-A-1 Première famille d'IRES :	32
	VII-2-A-2 Deuxième famille d'IRES :	33
	VII-2-A-3 Troisième famille d'IRES:	34
	VII-2-B Les facteurs cellulaires interagissant avec les IRES viraux	35
	VII-2-B-1 Les facteurs non canoniques :	35
	VII-2-B-2 IRES indépendantes des facteurs d'initiation canoniques :	36
	VII-3 Définition des types d'IRES	36
VIII	L'IRES du VHC : étude structurale et fonctionnelle	38
	VIII-1 Architecture de l'ARN virale	38
	VIII-2 Les domaines de l'IRES du VHC	39
	VIII-3 L'IRES du VHC est les facteurs de l'initiation CAP indépendante	40
IX	Le complexe eIF3/IRES	41
X	Objectifs de thèse	43

MATÉRIEL ET MÉTHODES	45
I Etude des protéines du complexe eIF3 humain	47
I-1 Biologie moléculaire	49
I-1-A Clonage des protéines : eIF3b, eIF3J, MRR-eIF3b et MRR-eIF3g	49
I-1-B Clonage par la technique LIC (Ligation Independent Cloning) des protéines du complexe eIF3.	49
I-2 Biochimie des protéines	51
I-2-A Expression et purification : eIF3b-MRR, eIF3g-MRR, eIF3j	51
I-2-B Expression et purification du domaine eIF3b-MRR pour l'étude RMN	52
I-2-C Expression et purification de la protéine eIF3b	52
I-2-D Purification du complexe eIF3	54
I-3 Méthodes de caractérisation	55
I-3-A Spectroscopie d'absorbance	55
I-3-B Dichroïsme circulaire	56
I-3-C Spectroscopie de fluorescence des résidus tryptophanes	56
I-3-D Diffusion de lumière multi-angles couplé à la chromatographie d'exclusion de taille et à la réfractométrie (SEC-MALLS-RI)	56
I-3-E Dynamic light scattering (DLS)	57
II Etude de l'IRES du VHC	57
II-1 Clonage des différents fragments d'IRES	57
II-2 Biochimie des ARN	58
II-2-A Préparation de la matrice ADN	58
II-2-B Protocole de transcription in vitro	58
II-2-C Obtention d'ARN radioactif	58
II-3 Purification	59
II-3-A Petite quantité d'ARN	59
II-3-B Grande quantité d'ARN	59
III Etudes d'interaction de complexe RNP	60
III-1 Activation de l'ARN	60
III-2 Filter Binding Assay	60
III-3 Retard sur gel «EMSA»	60
III-4 Etude RMN	61
IV Small angle X-Ray Scattering SAXS	61
IV-1 Protéine en solution	61
IV-2 ARN en solution	61
IV-3 Complexe Ribonucléoprotéique en solution	62
RÉSULTATS & DISCUSSIONS	
CHAPITRE I	65
Introduction	67
I Dissection moléculaire de l'IRES et obtention des échantillons.	70
I-1 Clonage, production et purification	70
I-2 Caractérisation	71
II La technique de SAXS	72
II-1 Principe du SAXS	72

II-2	Les différents outils informatiques	74
II-3	Principe de la construction de modèles	74
III	Modélisation des domaines individualisés	75
III-1	Technique de modélisation	75
III-2	Présentation des modèles	75
IV	SAXS versus RMN: étude du domaine II	78
V	Construction du modèle de l'IRES entier	79
VI	Comparaison de l'enveloppe de l'IRES avec les structures connues	83
VII	Etude de la dynamique du domaine II	87
VIII	Conclusion	89
CHAPITRE II		91
Introduction		93
I	Obtention de la protéine eIF3b humaine sous forme recombinante	96
II	Caractérisation de la protéine eIF3b et approche structurale	98
III	Etude structurale de la protéine eIF3b par SAXS	101
III-1	Etude du domaine MRR	101
III-2	Présentation de la structure du domaine MRR.	104
III-3	Réalisation de l'enveloppe du domaine MRR par SAXS	107
IV	Conclusion	113
CHAPITRE III		115
Introduction		117
I	La protéine eIF3b interagit-elle avec l'IRES du VHC?	118
II	Le domaine MRR de la protéine eIF3b interagit-il avec l'IRES du VHC?	120
III	Quels sont les acides aminés du domaine MRR-eIF3b impliqués ?	123
IV	Identification du domaine minimum de l'IRES du VHC impliqué dans l'interaction	125
V	Etude structurale du complexe MRR/IRES IIIabcd par SAXS	129
VI	Conclusion	131
CONCLUSION & PERSPECTIVES		137
EPILOGUE		143
PUBLICATION		147
RÉFÉRENCES BIBLIOGRAPHIQUES		163
ANNEXES		181
	Annexe I	182
	Annexe II	184
	Annexe III	188
	Annexe IV	189
	Annexe V	190
SUPPLÉMENTS		195
RÉSUMÉ		

TABLE DES ILLUSTRATIONS

Liste des figures

Titre	PAGE
FIG1 : Visualisation par microscopie des différents virus hépatiques.	9
FIG2 : Distribution de la prévalence mondiale du VHC.	12
FIG3 : Diagramme de l'épidémiologie du VHC.	13
FIG4 : Arbre phylogénique de la famille des <i>Flaviviridae</i> .	15
FIG5 : Les différents géotypes du VHC.	16
FIG6 : Reconstructions en 3D de particules VHC VLP non marquées (Yu et al 2007).	16
FIG7 : Présentation détaillée des constituants de la polyprotéine du VHC.	17
FIG8 : Cycle de vie du VHC.	19
FIG9 : Organisation de l'initiation de la traduction coiffe dépendante.	22
FIG10 : Déroulement de l'initiation de la traduction coiffe indépendante (cas du VHC) en comparaison de la traduction coiffe dépendante.	26
FIG11 : Analyse des sous-unités du complexe eIF3.	27
FIG12 : Présentation des données structurales du complexe eIF3.	29
FIG13 : Représentation schématique des IRES de la famille des Picornavirus (D'après la prédiction de structure secondaire).	33
FIG14 : Représentation schématique des IRES de la famille des <i>Flaviviridae</i> .	34
FIG15 : Représentation schématique de la famille des <i>Dicistrovirus</i> .	34
FIG16 : Diagramme simplifié représentant les 4 types d'IRES ainsi que leurs besoin en facteurs (canoniques et non canoniques) dans la voie de l'initiation de la traduction CAP indépendante.	37
FIG17 : Information structurale de l'IRES et présentation des zones d'interactions	39
FIG18 : Modèle en 3D de l'initiation de la traduction dans le cas du VHC par l'IRES conduisant à la formation du complexe 80S.	41
FIG19 : Diagramme des différentes approches techniques utilisées pour l'étude structurale et fonctionnelle du complexe eIF3/IRES-HCV.	43
FIG20 : Présentation des principales structures connues de l'IRES du VHC.	67
FIG21 : Dissection moléculaire de l'IRES du VHC.	69
FIG22 : Clonage, transcription et purification des différents domaines de l'IRES du VHC.	70
FIG23 : Résultats de Dynamic Light Scattering et vérification de l'homogénéité des échantillons d'ARN.	71
FIG24 : Principe du SAXS, exemple de la ligne ID14.3 de l'ESRF de Grenoble.	73
FIG25 : Dissection moléculaire de l'IRES du VHC et présentation des propriétés biophysiques.	77
FIG26 : Comparaison des données haute résolution et basse résolution sur le domaine II de l'IRES.	78
FIG27 : Remplacement moléculaire du domaine IIIabc par le programme CREDO et insertion des séquences hautes résolutions dans l'enveloppe de SAXS.	80

Titre	PAGE
FIG28 : Modélisation manuelle du domaine IIIabcdef.	81
FIG29 : Modélisation manuelle du domaine II-III.	81
FIG30 : Modélisation manuelle du domaine III-IV.	82
FIG31 : Modélisation manuelle de l'IRES II-III-IV par reconstruction à l'aide du programme VMD.	82
FIG32 : Superposition de la structure à 15Å dans l'enveloppe de l'IRES II-III-IV modélisée par SAXS.	83
FIG33 : Etude de l'IRES par microscopie électronique et construction d'un modèle 2D.	84
FIG34 : Présentation de l'IRES en complexe avec la sous-unité 40S.	85
FIG35 : Comparaison des modèles de l'IRES, libre ou en solution.	86
FIG36 : Modélisation de l'étude menée sur la dynamique du domaine II.	88
FIG37 : Modèle globale de l'activation de l'IRES du VHC lors de sa liaison à la sous-unité 40S.	89
FIG38 : Présentation des interactions de la protéine eIF3b dans le facteur eIF3.	94
FIG39 : Présentation de la protéine eIF3b.	94
FIG40 : Etude bioinformatique de la protéine eIF3b et identification de sous domaines.	95
FIG41 : Matrice de stabilisation.	96
FIG42 : Courbe théorique et étude de stabilité de la protéine eIF3 par thermal shift assay (TSA).	97
FIG43 : Purification de la protéine eIF3b sous forme recombinante.	97
FIG44 : Analyse de la structure secondaire de la protéine eIF3b par dichroïsme circulaire (CD).	99
FIG45 : Fluorescence des tryptophanes sur eIF3b.	100
FIG46 : Purification du domaine MRR-eIF3b.	101
FIG47 : Caractérisation de la protéine MRR-eIF3b.	102
FIG48 : Résultats de cristallogénèse et cristallographie de la protéine MRR-eIF3b.	103
FIG49 : Présentation du motif MRR de la sous-unité eIF3b.	105
FIG50 : Présentation des données biophysiques de la protéine MRR-eIF3b déterminées par SAXS et modalisation de l'enveloppe de la protéine.	106
FIG51 : Présentation des données biophysiques de la protéine eIF3b déterminées par SAXS et modélisation de l'enveloppe de la protéine.	108
FIG52 : Analyse bioinformatique de la protéine eIF3b par le serveur LOMETS.	110
FIG53 : Modélisation de la protéine eIF3b par insertion de séquence haute résolution dans l'enveloppe de SAXS.	111
FIG54 : Modélisation de la protéine eIF3b dans le complexe eIF3.	112
FIG55 : Présentation du complexe eIF3.	117
FIG56 : Etude de l'interaction du complexe eIF3 et de la sous-unité eIF3b sur l'IRES du VHC.	119

Titre	PAGE
FIG57 : Etude de l'interaction des domaines MRR avec l'IRES du VHC par FBA et EMSA.	121
FIG58 : Etude du complexe MRR-eIF3b / IRES, test de compétition et inhibition.	122
FIG59 : Résultats de l'étude de spectroscopie RMN menée sur l'identification des acides aminés du domaine MRR-eIF3b impliqués dans l'interaction avec l'IRES.	124
FIG60 : Identification du rôle du domaine III de l'IRES dans l'interaction avec le domaine MRR-eIF3b.	124
FIG61 : Illustration de la dissection moléculaire réalisée sur l'IRES du VHC.	126
FIG62 : Etude de l'interaction de l'IRES avec le motif MRR-eIF3b.	126
FIG63 : Résultats de l'étude de spectroscopie RMN menée sur l'identification des acides aminés du domaine MRR-eIF3b impliqués dans l'interaction avec l'IRES IIIId.	127
FIG64 : Etude par SAXS de l'interaction entre le domaine IIIId et le motif MRR-eIF3b.	128
FIG65 : Comparaison des résidus du MRR-eIF3b impliqués dans l'interaction avec la protéine eIF3j ou l'IRES du VHC.	130
FIG66 : Purification de la protéine eIF3j sous forme recombinante.	132
FIG67 : Formation du complexe eIF3 Δ j.	132
FIG68 : Formation du complexe eIF3 Δ j/eIF3jrec.	133
FIG69 : Mise en évidence expérimentale de la formation d'un complexe eIF3 Δ j/IRES en présence eIF3jrec.	134
FIG70 : Recrutement de la sous-unité 40S et du facteurs eIF3 par l'IRES du VHC.	141

Liste des tableaux

Titre	PAGE
Tableau I : les hépatites virales et les virus associés	7
Tableau II : Liste non exhaustive des différents traitements anti-VHC	14
Tableau III : Récapitulatif des facteurs nécessaires à l'initiation de la traduction	26
Tableau IV : Récapitulatif des sous-unité du complexe eIF3 et de leurs interactions	28
Tableau V : Liste non exhaustive des virus possédant un IRES	30
Tableau VI : Liste non exhaustive des différents IRES cellulaires	31
Tableau VII : Liste des constructions réalisées pour l'étude	47
Tableau VIII : Liste des vecteurs utilisés pour le clonage et l'expression des gènes	48
Tableau IX : Liste des anticorps utilisés pour l'immunodetection	55
Tableau X : Protocole de la transcription <i>in vitro</i>	59
Tableau XI : Liste des constructions réalisées pour l'étude de SAXS	68
Tableau XII : Liste des différents paramètres biophysiques obtenus par SAXS	72
Tableau XIII : Fonction des différents programmes	74
Tableau XIV : Présentation des domaines de l'IRES générés par biologie moléculaire et de leur affinité pour MRR-eIF3b.	125

ABREVIATIONS

Å : Angström
aa : acide aminé
ADN : Acide Désoxyribonucléique
AI : Auto-inductible
ARN : Acide Ribonucléique
ARNm : Acide Ribonucléique messenger
BUNCH : Modelling of Multidomain Proteins Against Multiple Data Sets
CAP : Coiffe
cpm : coup par minute
CREDO : Add Missing Loops and Domains to High and Low Resolution Models of Proteins
CRYSOL : Evaluation of X-ray Solution Scattering Curves from Atomic Models
DAMMIF : Rapid ab initio shape determination by simulated annealing
DAMMIN : Ab initio shape determination by simulated annealing
DLS : Dynamic light scattering
DMSO : diméthylsulfoxyde
DTT : Dithiothreitol
eIF : eukaryotic initiation factor
EMSA : Electro-Mobility-Shift-Assay
EOM : Ensemble Optimization Method
ESRF : European Synchrotron Radiation Facility
FBA : Filter Binding Assay
GASBOR : Ab initio Reconstruction of a Protein Structure by a Chain-Like Ensemble
GNOM : Small-Angle Scattering Data Processing Using the Regularisation Technique
GST : Glutathion S transferase
HNANB : Hépatite Non A non B
IPTG : IsoPropyl-beta-D-ThioGalactopyranoside
IRES : Internal Ribosome Entry Site
kb : kilobase
Kd : Constante de dissociation
kDa : kiloDalton
LB : Luria Broth
LIC : Ligation Independant Cloning
MALLS-SEC : Multiple Angle Laser Light Scattering Size Exclusion Chromatography
MBP : Maltose Binding Protein
MRR : Motif de Reconnaissance à l'ARN
NSD : Normalized Spatial Discrepancy
nt : Nucléotides
ORF : Open Reading fram
PAGE : Polyacrylamide Gel Electrophoresis
PCR : Polymérisation en chaîne
pdb : Protein data bank
PRIMUS : Primary Analysis & Manipulations with Small-Angle Scattering Data
RE : Réticulum Endoplasmique
RNP : Ribo-Nucleo-Protéin
Rg : Rayon de gyration
RMN : Résonance Magnétique Nucléaire
SANS : Small Angle Neutron Scattering
SASREF : Modelling of multisubunit complexes
SAXS : Small Angle X ray Scattering
SW : Swing (type de rotor)
TEV : Tabacco etch virus
UTR : Untranslated Region
VH(A, B, C, D, E, G) : Virus de l'Hépatite

VIH : Virus de l'immunodéficience Humaine

VLDL : Very Low Density Lipoprotein

VLP : Lipo-Viro-Particle

Code à une et trois lettres pour les 20 acides aminés :

G - Glycine (Gly)	P - Proline (Pro)
A - Alanine (Ala)	V - Valine (Val)
L - Leucine (Leu)	I - Isoleucine (Ile)
M - Methionine (Met)	C - Cysteine (Cys)
F - Phenylalanine (Phe)	Y - Tyrosine (Tyr)
W - Tryptophan (Trp)	H - Histidine (His)
K - Lysine (Lys)	R - Arginine (Arg)
Q - Glutamine (Gln)	N - Asparagine (Asn)
E - Glutamic Acid (Glu)	D - Aspartic Acid (Asp)
S - Serine (Ser)	T - Threonine (Thr)

PROLOGUE

L'homme a, depuis toujours, été sujet à analyser, chercher et comprendre ce qui l'entoure comme ce qui le compose. Depuis plus de 3000 ans, le foie est l'objet de nombreuses questions. C'est au V^{ème} siècle avant J.-C., que le père de la médecine actuelle, Hippocrate, fit le premier, le lien entre le foie et la présence d'une coloration jaune de la peau et des muqueuses en cas de trouble hépatique. C'est dans le Corpus Hippocratique (recueil d'une soixantaine de livres de médecine écrits en ionien) qu'est cité, pour la première fois, le terme «ictère», transformé par les latins en *icterus*, pour désigner la jaunisse. Les médecins de l'époque décrivaient les symptômes de personnes avec de la fièvre et des vomissements qui, dans les sept jours suivants, développaient une jaunisse, description pouvant correspondre à la présence d'hépatite virale. Ce n'est que plusieurs siècles après, que l'on découvrit que l'hépatite était une maladie transmissible, éventuellement par un agent infectieux, via les seringues et la transfusion sanguine. Ces faits ont été constatés et avérés au cours des nombreuses guerres du XVII^{ème} au XIX^{ème} siècles. Il fallut attendre la fin des années 1980 pour qu'un nouvel agent viral, responsable de la majorité des hépatites post-transfusionnelles non lié aux hépatites connus, fut mis en évidence par des techniques de biologie moléculaire. Pour la première fois de l'histoire de la virologie, un virus était identifié par son génome sans que la particule virale soit elle-même isolée. Ce virus fut nommé virus de l'hépatite C ou VHC. L'infection par le VHC est un problème de santé publique. De ce fait, il fait l'objet de nombreux travaux de recherche afin d'appréhender sa biologie moléculaire et son cycle viral. La nature ARN de son génome favorise une dérive génétique importante lors de son processus de réplication. Cette variabilité lui confère une capacité particulière à s'adapter à son hôte et ralentit de ce fait les traitements antiviraux. Cependant, ce génome à ARN possède une région en 5' UTR fortement conservée et impliquée dans l'initiation de la traduction des protéines virales. En dépit de l'importance du VHC et des nombreuses études portant sur celui-ci, beaucoup d'éléments restent encore à être élucidés sur cet ARN. Ce mémoire a comme sujet, l'étude des événements impliqués dans l'initiation de la traduction coiffe indépendante par l'IRES du VHC.

INTRODUCTION

Virus et hépatites

I Les hépatites virales

Une hépatite est une inflammation du foie d'origine toxique, auto-immune ou virale. En 1947, Mac Callum (Mac 1947) étudia pour la première fois des hépatites épidémiques avec la notion de transmission par voie sanguine. En 1966, l'équipe de Blumberg put faire la relation entre une atteinte hépatique et la présence d'antigène sérique (Blumberg *et al.* 1966). En 1970, DS Dane visualisa pour la première fois à l'aide d'un microscope électronique le Virus de l'Hépatite B (VHB) sous forme d'une sphère de 42 nm de diamètre (Dane *et al.* 1970). En 1973, S. Feinstone étudia l'hépatite A qui fut découverte chez des patients atteints d'hépatite (Feinstone *et al.* 1973). Les équipes de S. Feinstone et de H.J Alter signalèrent dès 1975 la présence de cas d'hépatites post-transfusionnelles en l'absence de marqueurs caractéristiques de l'infection par les virus A et B (Feinstone *et al.* 1975). Ils classifièrent ainsi pour la première fois les hépatites virales en deux groupes : A et B. Cette classification sera beaucoup plus tard étayée par l'identification de nouveaux symptômes et d'une nouvelle hépatite dénommée non A non B (HNANB). A ce jour, six virus ont été identifiés pouvant provoquer une inflammation du foie (**Tableau I**). Les HNANB regroupaient donc les hépatites C, D, E et G.

Tableau I : Les hépatites virales et les virus associés (*n.a = non attribué)

	VHA	VHB	VHC	VHD	VHE	VHG
Famille virale	<i>Picornaviridae</i>	<i>Hepadnaviridae</i>	<i>Flaviviridae</i>	Satellite du VHB	n.a	<i>Flaviviridae</i>
Genre	Hepatovirus	Orthohepadnavirus	Hepacivirus	Deltavirus	n.a	n.a
Symétrie	Icosaédrique	Icosaédrique	Icosaédrique ?	Sphérique	Icosaédrique	Icosaédrique
Enveloppe	Non	Oui	Oui	Oui	Non	n.a
Taille	27 à 32nm	42 à 47nm	≈ 60nm	36nm	30 à 34nm	65nm
Génome	ARN linéaire 7,5kb	ADN circulaire partiellement bi-caténaire de 3,2kb	ARN linéaire simple brin positif de 9,5kb	ARN circulaire simple brin négatif de 1,7kb	ADN circulaire partiellement bi-caténaire de 7,5kb	n.a*
Voie de transmission	Entérique	Sanguine, sexuelle, verticale	Sanguine	Sanguine, verticale	Entérique	Sanguine et sexuelle
Temps d'incubation	30 jours	30 à 120 jours	4 à 12 semaines	40 jours à 6 mois	15 à 60 jours	n.a
Présence de porteurs chroniques	Non	Oui	Oui	Oui	Non	Oui
Evolution en hépatite chronique	Non	Oui	Oui	Oui	Non	n.a
Evolution en hépatite fulminante	Oui	Oui	Rare	Oui	Oui	n.a
Traitement curatif	Non	Interféron- α	Interféron- α Ribavirine	Non	Non	Non
Vaccin	Oui	Oui	Non	Non	Non	Non

Tous ces virus disposent du même cycle d'infection malgré la présence d'un arsenal génétique très varié. Ces virus, après s'être introduits dans l'organisme, se dirigent préférentiellement vers le foie pour s'y multiplier. Les hépatocytes ainsi infectés vont activer le système immunitaire en provoquant une inflammation locale plus ou moins importante en fonction de l'état du patient et de la souche virale. C'est en grande partie l'hyperactivité du système immunitaire qui va provoquer la destruction des cellules infectées et provoquer ainsi une nécrose partielle du foie pouvant évoluer jusqu'à la cirrhose. Les symptômes caractéristiques d'une hépatite aiguë peuvent durer plusieurs semaines : jaunissement de la peau et des yeux, urines foncées, fatigue importante, nausées ou encore douleurs abdominales. Certains virus comme le Virus de l'Hépatite C (VHC) peuvent conduire dans 80% des cas à un état dit de « porteur chronique » qui signifie que le système immunitaire du sujet est globalement inefficace et ne parvient pas à éliminer le virus. Cet état peut évoluer des années plus tard vers un cancer du foie.

L'hépatite A est une infection virale du foie, due à la présence du Virus de l'Hépatite A (VHA). C'est un virus nu de la famille des *Picornaviridae* et du genre *Hepatovirus* avec une taille de 28 nm de diamètre, identifié en 1973 (**FIG1A, Tableau I**). Ce virus est composé d'une capsid de symétrie icosaédrique possédant un génome à ARN simple brin de 7,8 kb. La contamination (1,5 million de nouveaux cas par an dans le monde) se fait par contact avec une personne contaminée ou par l'ingestion d'aliments ou d'eaux contaminées (voie oro-fécale) (Nothdurft *et al.* 2008). Un vaccin très efficace existe depuis 1997 et fait toujours l'objet de campagne vaccinale en Europe comme dans les pays en voie de développement.

L'hépatite B est une infection virale du foie due à la présence du Virus de l'Hépatite B (VHB). Le virus associé appartient à la famille des *Hepadnaviridae* et au genre *Orthohepadnaviru*. Sa taille est de 42 nm de diamètre (**FIG1B, Tableau I**). Ce virus est composé d'une capsid de symétrie cubique. Son génome est composé d'un ADN partiellement double brin de 3200 nt. L'hépatite B est considérée comme une maladie infectieuse extrêmement contagieuse puisque le VHB est 100 fois plus infectieux que le Virus de l'Immunodéficience Humaine (VIH). L'hépatite B est l'une des principales maladies humaines : on estime à 2 milliards le nombre de personnes ayant été infectées par le virus. Parmi elles, 350 millions de personnes sont des porteurs chroniques (10% des sujets infectés) et sont exposés à un risque élevé de décès dû à un cancer du foie (un million de morts chaque année dans le monde). En raison de la faible efficacité des traitements antiviraux, la vaccination contre l'hépatite B est la principale mesure fiable pour se protéger contre cette maladie. Depuis 1981, on dispose de vaccins anti-hépatite B très efficaces (GenHevac B, Pasteur – Engerix GSK).

L'hépatite C est une infection virale du foie due à la présence du Virus de l'Hépatite C (VHC) (**FIG1C, Tableau I**). Ce virus appartient à famille des *Flaviviridae* et du genre *Hepacivirus* avec une taille de 60 nm de diamètre (Petit *et al.* 2005, Andre *et al.* 2006). Il a été identifié en 1974 comme responsable des HNANB. Son génome est composé d'un ARN de polarité positive avec une taille de 9,7kb. Il affecte 3% de la population mondiale et se transmet principalement par

voie sanguine. Il n'existe pas de vaccin, et seule une thérapie à base d'interféron- α et de ribavirine est actuellement disponible (**Tableau I**).

L'hépatite D est une infection virale du foie due à la présence du Virus de l'Hépatite D (VHD). Le VHD est un virus répertorié comme défectif et seulement identifié chez des patients infectés par le VHB. Il s'agit d'une particule sphérique de 36 nm de diamètre dont l'enveloppe est

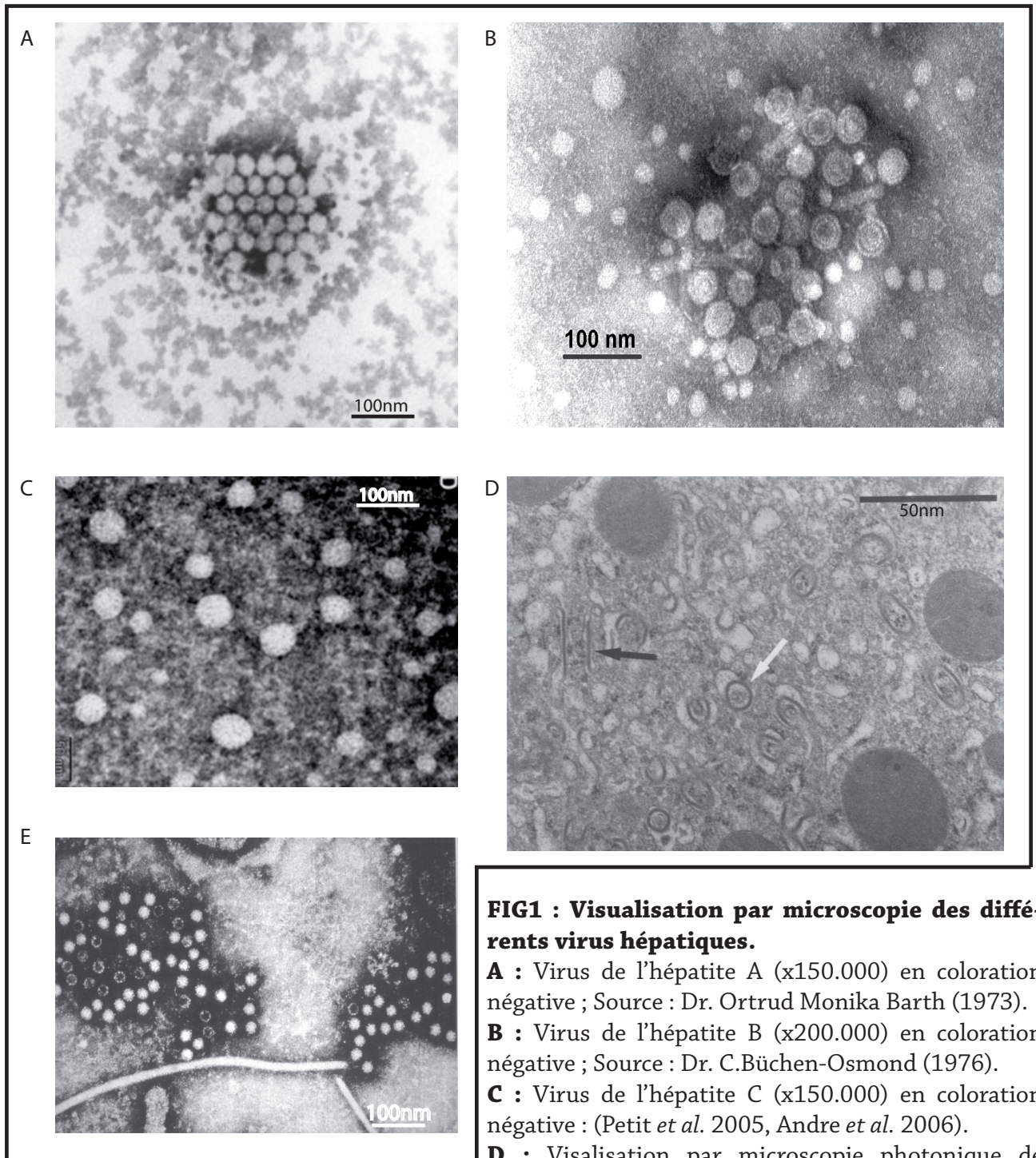


FIG1 : Visualisation par microscopie des différents virus hépatiques.

A : Virus de l'hépatite A (x150.000) en coloration négative ; Source : Dr. Ortrud Monika Barth (1973).

B : Virus de l'hépatite B (x200.000) en coloration négative ; Source : Dr. C.Büchen-Osmond (1976).

C : Virus de l'hépatite C (x150.000) en coloration négative : (Petit *et al.* 2005, Andre *et al.* 2006).

D : Visalisation par microscopie photonique de

tubules cytoplasmiques suite à l'infection par le VHC. Source : (Shimizu *et al.* 1979).

E : Virus de l'hépatite E (x150.000) en coloration négative ; Source : Centers for disease control.

constituée de l'enveloppe du VHB portant les déterminants de l'antigène HBs. Le génome est composé d'un ARN simple brin de 1,75 kb qui se rapproche par son organisation génétique de certains ARN pathogènes chez les plantes. Il est à noter que, contrairement au VHB, il n'existe pas de vaccin efficace, et que seule une thérapie à base d'interféron- α est prescrite (tableau I).

L'hépatite E est une infection virale du foie due à la présence du Virus de l'Hépatite E (VHE) (**FIG1E, Tableau I**). Il a été classé initialement dans la famille des *Caliciviridae*. Toutefois, son génome ressemble de très près à celui du virus de la rubéole, il est maintenant classé comme membre unique du genre *Hepevirus*, et aucune famille n'a encore été rattachée à ce genre (Tableau I). C'est un virus mesurant 30 nm de diamètre qui a été identifié en 1983. Ce virus ne comprend pas d'enveloppe et contient un brin d'ARN enroulé dans le sens positif long d'environ 7300 nt.

L'hépatite G est une infection virale du foie due à la présence du Virus de l'Hépatite G (VHG). Il s'agit d'un groupe de virus (GBV-A, B et C) appartenant à la famille des *Flaviviridae* avec une taille de 65 nm de diamètre, identifié en 1995. Le virus GBV-C est phylogénétiquement étroitement lié au VHC, mais semble se reproduire principalement dans les lymphocytes et les cellules monocytaires du sang périphérique, et peu ou pas du tout dans les hépatocytes (**Tableau I**).

II Présentation détaillée de l'hépatite C

II-1 Historique

C'est en 1974 que l'existence d'hépatites sérologiquement dissemblables des hépatites A et B a été reconnue (tableau I). Ces hépatites furent appelées provisoirement HNANB par l'équipe de Feinstone (Feinstone, Kapikian *et al.* 1975) et représentaient environ 90 % des hépatites post-transfusionnelles. Alter et son équipe montrèrent en 1978 que le sérum de quatre patients atteints de HNANB pouvait infecter des chimpanzés, démontrant ainsi qu'au moins un des agents des HNANB était un agent transmissible (Alter *et al.* 1978). Pendant plusieurs années, le seul moyen d'étude de cet agent résidait dans l'utilisation du modèle animal « chimpanzé », ce qui permit à de nombreuses équipes d'étudier et de caractériser les conséquences cytologiques de l'infection par la formation de tubules cytoplasmiques (**FIG1D**) inobservables chez l'humain (Shimizu *et al.* 1979). En 1981, l'équipe de Shimizu et l'équipe d'Hollinger ont montré qu'au moins deux agents distincts étaient responsables des symptômes provoqués par une NANBH. En 1983, les travaux de Feinstone montrèrent que le chloroforme ôtait le pouvoir infectieux de l'agent responsable des tubules cytoplasmiques (Feinstone *et al.* 1983), suggérant que cet agent était potentiellement un virus enveloppé. De plus, des expériences de filtration permirent de montrer que ce virus était de petite taille. Cependant la définition réelle du virus n'était pas encore clairement établie. C'est seulement que plus tard, en 1989 que le génome du virus fut caractérisé par l'équipe de M. Houghton (Choo *et al.* 1989). Ils utilisèrent du sérum de chimpanzé infecté par l'agent provoquant des HNANB pour produire une large banque de

cDNA qui fut ensuite clonée dans un bactériophage. Cette banque d'expression fut alors criblée par des sérums de patients humains atteints de NANBH et porteurs d'anticorps spécifiques. Les auteurs isolèrent le clone 5.1.1, codant un antigène reconnu par les sérums de plusieurs patients. De plus, cette équipe réussit à caractériser le génome du virus, il s'agissait d'un ARN simple-brin codant pour un unique ORF. L'agent causal des HNANB, ainsi identifié, prit alors le nom de VHC. D'un point de vue médical, cette découverte permit de mettre au point un test ELISA pour détecter la présence d'anticorps spécifiques du VHC chez les patients. De plus, ce test montra que le VHC était la principale cause des HNANB transfusionnelles à travers le monde (Kuo *et al.* 1989) et permit d'éliminer la première cause de transmission du virus dans les pays développés. Une fois le génome du virus caractérisé, de nombreuses études ont permis de préciser la structure secondaire et tertiaire des extrémités 5' et 3' non traduites du génome ou encore de caractériser le mécanisme de clivage de l'unique polyprotéine, ainsi que les différentes fonctions des protéines virales. En 1992, l'équipe de A. Nomoto a montré que l'extrémité 5' non traduite contenait une région structurée hautement conservée, correspondant à un site d'entrée interne des ribosomes (IRES) (Tsukiyama-Kohara *et al.* 1992). En 1993, l'équipe de Bartenschlager montra que l'extrémité N-terminale de la protéine NS3 possédait une activité serine protéase en comparant le génome de HCV à celui des *Pestivirus* (Bartenschlager *et al.* 1993). Le rôle hélicase de cette protéine ne fut démontré qu'en 1996 par l'équipe de Tai (Tai *et al.* 1996). En 1996, l'équipe de Behrens a ensuite montré que la polyprotéine possédait une activité ARN-polymérase ARN-dépendante portée par la protéine NS5B (Behrens *et al.* 1996). Néanmoins, les recherches sur le VHC se trouvaient fortement limitées par l'absence de modèle cellulaire, car seul le chimpanzé pouvait être infecté et reproduire une infection et une hépatite proche de celle observée chez l'homme. Une avancée majeure fut la mise au point en 1999 par l'équipe de Ralf Bartenschlager du premier réplicon dérivé du génome de HCV (Lohmann *et al.* 1999). Il aura fallu attendre 2002 pour que l'équipe de V. Lotteau (Lyon) caractérise une nouvelle forme du virus, jouant un rôle important dans la dissémination virale *in vivo* : des VLP (Viro-lipido-particles) associant particules virales et VLDL (Very-Low-Density-Lipoprotein) (Andre *et al.* 2002). En 2004, deux découvertes sont venues compléter nos connaissances sur le spectre d'infection de ce virus. En effet, le virus, bien que majoritairement présent dans le foie, a été trouvé dans le système nerveux central (Forton *et al.* 2004) et dans l'intestin, où l'infection semble être également productrice de particules virales (Deforges *et al.* 2004). En 2003, l'équipe de Charles Rice décrivit un système permettant pour la première fois de reproduire l'intégralité du cycle viral en culture cellulaire (Kato *et al.* 2003, Wakita *et al.* 2005, Lindenbach *et al.* 2005, Zhong *et al.* 2005) ouvrant une nouvelle voie pour les recherches fondamentales sur ce virus. La première image de microscopie électronique du VHC a été réalisée en 2005 par l'équipe de Christian Trépo (Lyon) (**FIG1C**).

II-2 Epidémiologie

On estime que le VHC affecte plus de 200 millions de personnes dans le monde, soit environ 3% de la population, ce qui est quatre fois plus que la population atteinte par le VIH. De plus, 3 à 4 millions de personnes sont nouvellement infectées chaque année. Les pays les plus touchés sont l'Égypte, le Burundi, le Cameroun, la Guinée, la Bolivie, la Tanzanie et la Mongolie, et en Europe la Roumanie (**FIG2**). Si on regarde de près la situation en Egypte (plus de 10% de sujets infectés), on suppose que le nombre de patients infectés a effectué un bon en avant suite à la campagne de vaccination contre la bilharziose, menée en Egypte dans les années 1980 jusqu'en 1990. En France, selon une étude de 2004, 780 000 personnes seraient infectées par le VHC. On considère que ce virus est responsable d'environ 10 % des cas d'hépatites aiguës et de 80 % des cas d'hépatites chroniques (**FIG3**). L'hépatite chronique C est une cause majeure de cirrhose et de cancer primitif du foie (carcinome hépato-cellulaire). L'évolution silencieuse de la maladie et la fréquence élevée de passage à la chronicité explique l'existence d'un grand réservoir de sujets infectés, dont une grande partie ignore son état de séropositivité.

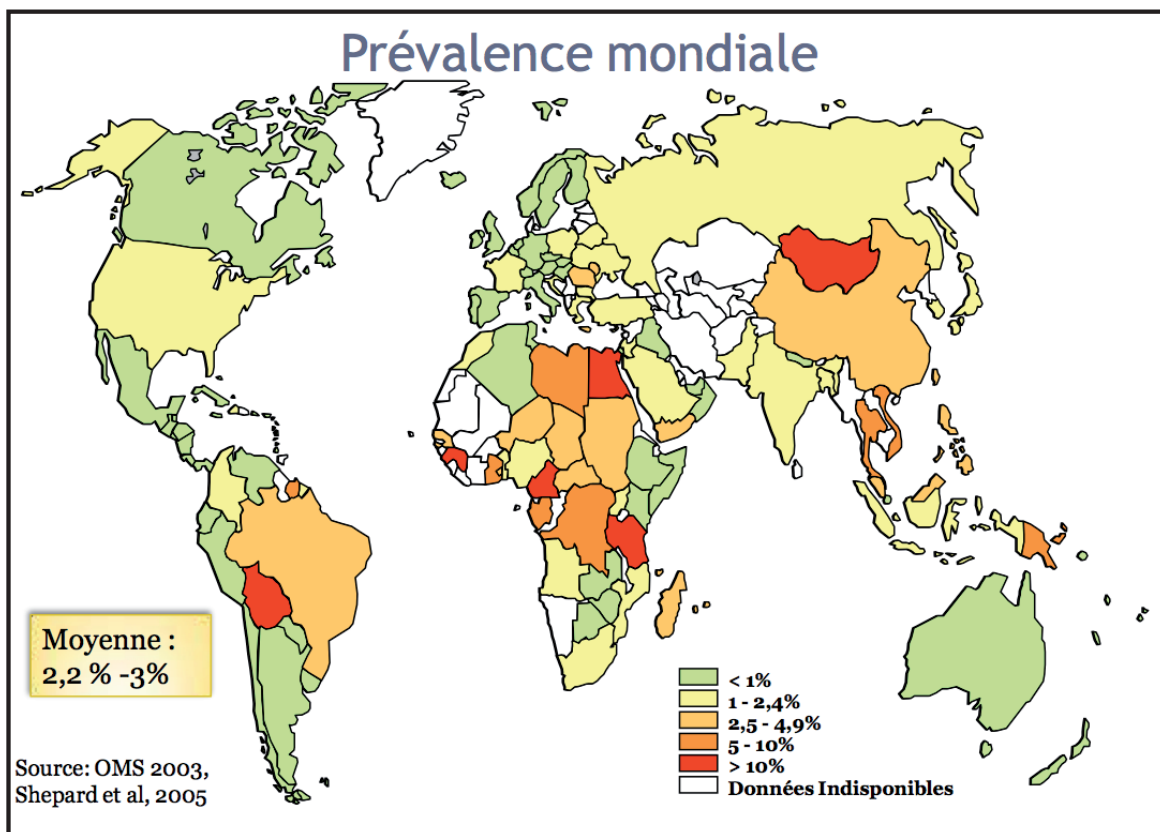


FIG2 : Distribution de la prévalence mondiale du VHC.

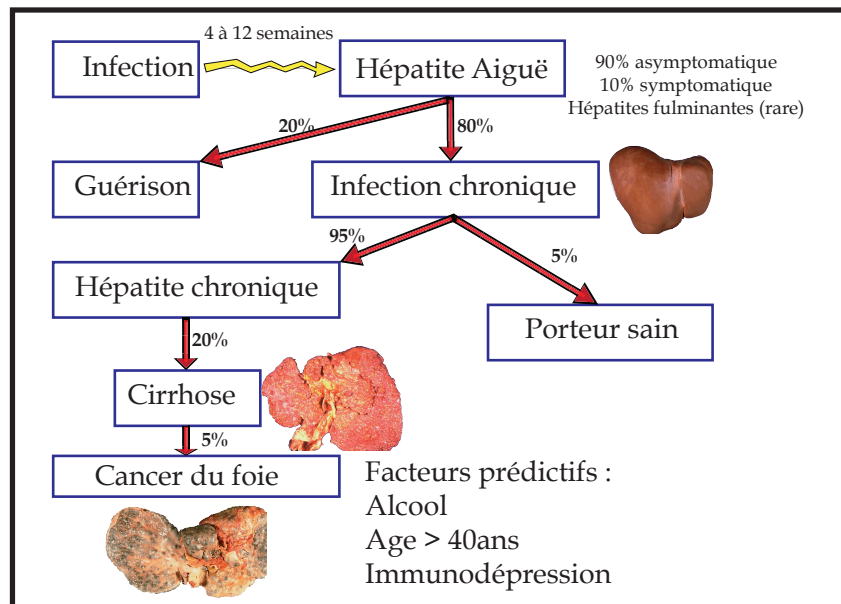


FIG3 : Diagramme de l'épidémiologie du VHC.

II-3 Transmission

Il y a 20 ans, le VHC se transmettait principalement par voie sanguine suite aux transfusions et actes chirurgicaux. Après la découverte de la présence du virus dans les produits sanguins, la mise en place du dépistage systématique de la présence des anticorps anti-VHC a permis de diminuer très fortement la transmission post-transfusionnelle. De plus, d'importants moyens de décontamination ont permis de diminuer les cas de transmission à la suite d'actes chirurgicaux (dentiste, opérations). Aujourd'hui, 70% des transmissions sont causées par l'utilisation de systèmes d'injections contaminés (toxicomanie et tatouage). D'après une étude publiée en 2007 par l'équipe de A.J. Tatem (Pybus *et al.* 2007) il serait possible que la dissémination du VHC impliquerait également d'autres mécanismes. Effectivement, ces auteurs ont constaté la présence d'un autre mode de transmission ubiquitaire impliquant des arthropodes. La présence de cette voie supplémentaire pourrait expliquer la dissémination de certains géotypes du VHC à travers le monde. Cependant, il s'agirait d'un mécanisme peu infectant, mais suffisant pour maintenir un niveau basal de sujets infectés dans la population. De plus, cet hôte pourrait jouer un rôle dans la divergence génotypique du VHC. Cette hypothèse demande à être confirmée par des résultats expérimentaux, mais elle indique la présence d'une voie « naturelle » de transmission du VHC.

II-4 Maladie et traitements

L'hépatite C aiguë est asymptomatique, ce qui explique la difficulté du diagnostic au stade aigu de la maladie. Dans les formes symptomatiques, la maladie dure généralement de 2 à 12 semaines. Les symptômes ne sont pas spécifiques et souvent dissimulés par l'anorexie, fatigue, nausées, douleurs suivies de l'apparition d'urines foncées et d'une jaunisse dans 25% des cas. Ils sont semblables à ceux observés au cours d'autres hépatites virales ce qui rend le

diagnostic plus complexe. Dans environ 80% des cas, l'infection évolue vers la chronicité qui est elle-même asymptomatique ; son diagnostic est donc occasionnel. On estime que 20% des personnes chroniquement infectées par le virus de l'hépatite C développeront une cirrhose hépatique dans les 20 premières années qui suivent la contamination.

Les formes aiguës évoluent d'elles même favorablement dans la majorité des cas sans nécessiter de prescription médicamenteuse, mais une guérison complète spontanée de l'hépatite aiguë C n'est observée que dans 30% des cas. Les formes chroniques responsables de fibroses hépatiques importantes doivent être traitées par l'association interféron-pégylé et ribavirine (Layden-Almer *et al.* 2003; Feld et Liang 2005; Chevaliez et Pawlotsky 2008). De plus, ce traitement (onéreux) comporte de très nombreux effets indésirables avec, entre autres, des symptômes neuropsychiatriques, des symptômes pseudo-grippaux ou encore des anomalies hématologiques. La thérapie avec l'interféron- α (combiné au polyéthylène glycol) administré en bithérapie avec de la ribavirine est devenue le traitement standard pour les personnes atteintes d'hépatite C chronique. Cependant, le taux de réponse au traitement chez les patients n'est que de 55 % (cas du génotype 1b) et les effets secondaires constituent un obstacle important (Manns et Strassburg 2001; Fried 2002; Wiegand *et al.* 2006). En conclusion, les thérapies contre le VHC sont globalement insatisfaisantes (Toyoda *et al.* 2000; Koshy *et al.* 2002), il devient donc urgent d'identifier de nouvelles cibles, permettant la synthèse d'anti-viraux, qui préviendraient la propagation de la maladie (**Tableau II**). Une liste plus détaillée est présentée en **Annexe I**, (page 182).

Tableau II : Liste non exhaustive des différents traitements anti-VHC (source : 2008)

Cible	Mécanisme	Compagnie pharmaceutique	Phase clinique
IRES	ARN Antisens	ISIS Pharmaceuticals	Phase II
	Ribozyme	Elan corporation	Phase III
		Ribozyme Pharmaceuticals	
NS3	Inhibiteur d'une sérine protéase	Boehringer ingelheim	Phase III
	Inhibiteur d'une sérine protéase	Vertex Pharmaceuticals	Pré-clinique
NS5B	Inhibiteur RdRp	Japan tobacco	Phase II
E1	Vaccin thérapeutique	innogenetics	Phase II
E2	Anticorps monoclonal	XTL BioPharmaceuticals	Phase
IMPDH	Inhibiteur	Vertex Pharmaceuticals	Phase III
		Ribapharm	Phase I
		Ribapharm	Phase I
Fibrose du foie	Anti-fibrotique	Intermune Pharmaceuticals	Phase II
		Intermune Pharmaceuticals	Phase III
Apoptose du foie	Inhibiteur des caspases	Idun Pharmaceuticals	Phase II
HCC	Inhibiteur de la b-tubuline	Tularik	Phase III
Ré-infection du VHC	Immunoglobuline anti HCV avec activité immunosuppresseur	Nabi Pharmaceuticals	Phase II
		Roche Holdings	Pre-clinique

A l'heure actuelle, il n'existe aucun vaccin contre l'hépatite C. Le moyen le plus efficace de lutter contre l'hépatite C est de diminuer le risque de transmission par des systèmes d'injections

contaminés, et de réajuster les comportements de la population à risque. Au Québec, dans la ville de Montréal, il y a depuis quelques années une recrudescence du VHC dans le quartier sud de la ville. Les autorités locales n'ont eu, pour seule arme contre ce virus, que la distribution de seringues stériles pour les toxicomanes.

III Le virus de l'hépatite C

III-1 Classification

Le VHC fait partie de la famille des *Flaviviridae*, comprenant les genres *Flavivirus*, *Pestivirus* et *Hepacivirus* (**FIG4**). Les *Flaviviridae* incluent plusieurs pathogènes humains importants, dont les virus de la dengue (DV), de la fièvre jaune (YF), de l'encéphalite transmise par les tiques (TBE) et celui du Nil occidental (WNV). Ces virus partagent certaines caractéristiques physiologiques dont un génome à ARN simple brin de polarité positive, d'une longueur de 9500 à 12 500 nucléotides ; Le VHC est le seul représentant du genre *Hepacivirus*.

Son génome diffère considérablement entre les 6 génotypes (**FIG5**) et regroupe plus de 90 sous-types. En effet, on note qu'il peut y avoir jusqu'à 34 % de divergence entre les séquences nucléotidiques de 2 génotypes du VHC (Simmonds *et al.* 2005).

III-2 Structure du VHC

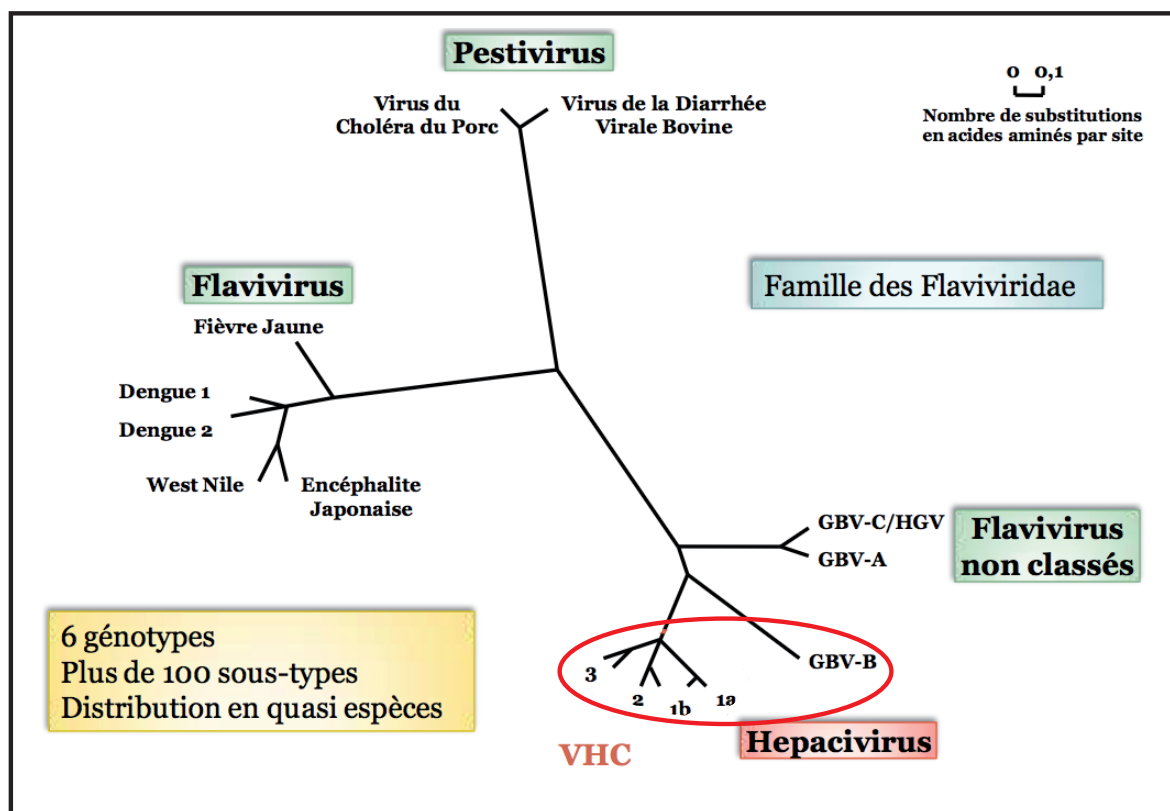


FIG4 : Arbre phylogénétique de la famille des Flaviviridae (Simons *et al.* 2005).

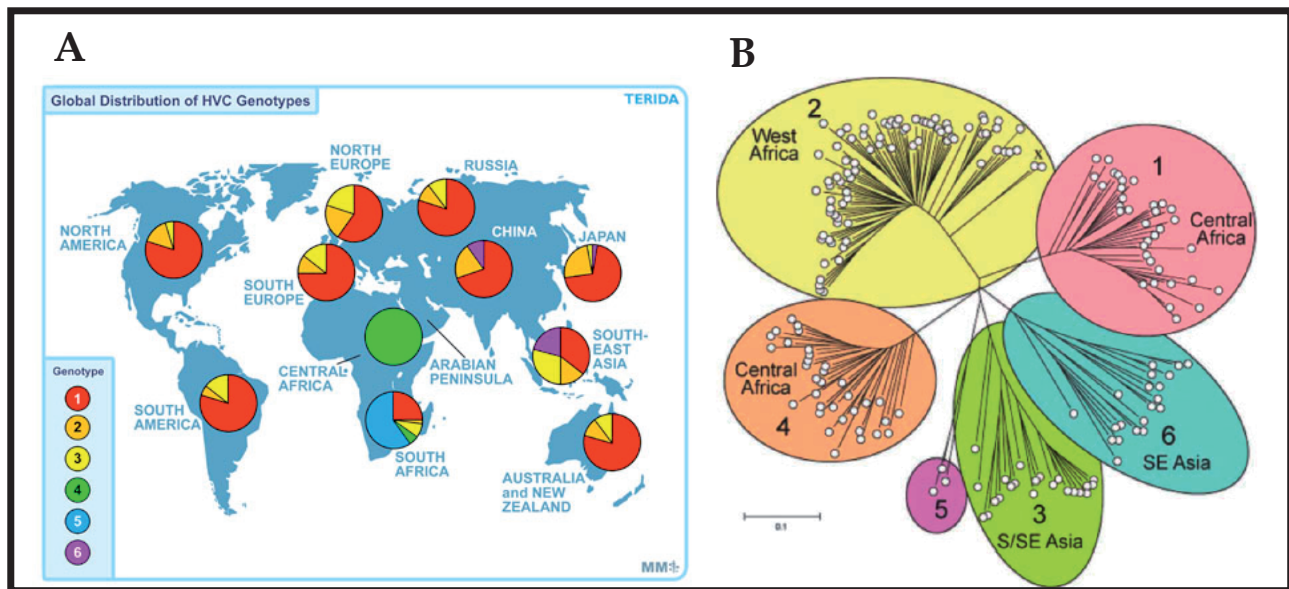


FIG5 : Les différents génotypes du VHC

A : Distribution mondiale des différents génotypes du VHC.

Remarque : En France le génotype 1b est prédominant, (Zein *et al.* 2000).

B : Arbre phylogénique des différents génotypes, basé sur la conservation de séquences de la protéine NS5B (Simmons *et al.* 2005).

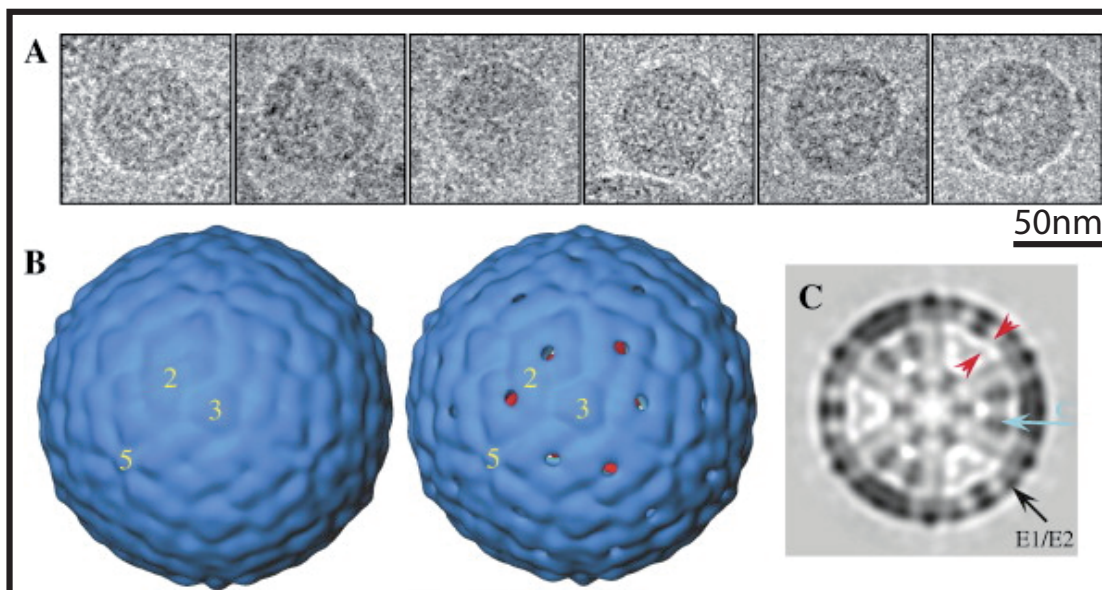


FIG6 : Reconstructions en 3D de particules VHC VLP non marquées (Yu et al 2007).

A : Clichés de cryo-electromicroscopie

B : Représentation en 3D de la surface des particules avec une résolution de 30Å.

C : Reconstruction de la membrane des VLP à 7Å de résolution, indiquant la présence d'une membrane en multi-couche.

E1 et E2 sont les protéines de surface.

L'ARN simple brin est logé dans la nucléocapside (Takahashi *et al.* 1992; Ishida *et al.* 2001; Maillard *et al.* 2001) entourée d'une enveloppe lipidique, provenant du réticulum endoplasmique (RE) de la cellule infectée, sur laquelle sont ancrées deux glycoprotéines virales, E1 et E2 (Cocquerel *et al.* 1999). Des particules provenant de tissus de foie et de sérums infectés ont pu être visualisées directement par microscopie électronique (ME) (Kaito *et al.* 1994; Baumert *et al.* 1998). Ces particules sont hétérogènes et ont des diamètres variant de 55 à 65 nm (Maillard, Krawczynski *et al.* 2001; Yu *et al.* 2007) (**FIG6**).

III-3 Organisation du génome et maturation de la polyprotéine

Le génome du VHC possède un seul et unique cadre de lecture ouvert codant pour une polyprotéine d'environ 3 000 acides aminés. L'ARN est bordé, en 5' et en 3', de régions non codantes (5' et 3' UTR) d'environ 340 et 230 nt respectivement (Grakoui *et al.* 1993a, b, c). L'extrémité 5'-UTR est une région hautement conservée repliée en quatre domaines majeurs nommés de I à IV (Brown *et al.* 1992; Honda *et al.* 1996; Honda *et al.* 1996; Honda *et al.* 1999). Après leur traduction, les protéines structurales sont clivées par des peptidases cellulaires, alors que les protéines non structurales sont libérées par des protéases virales. La région d'ARN codant pour la polyprotéine est divisée en 2 régions : une partie N-terminale qui code pour les protéines structurales (C et E1 et E2) alors que les deux tiers de l'ARN restants codent pour les protéines non structurales (**FIG7**).

La protéine **C** (Core), est un composant essentiel de la capsid. Elle est fortement basique avec une taille de 22kDa. Cette protéine est impliquée dans de nombreuses fonctions cellulaires, telles que l'apoptose, la transformation cellulaire ou encore la signalisation intracellulaire du métabolisme lipidique (Kato *et al.* 2001).

E1 et **E2** sont des glycoprotéines d'enveloppe qui s'associent en hétéro-dimères avec des tailles de 30 et 70 kDa respectivement (Hijikata *et al.* 1991b; Selby *et al.* 1993; Grakoui *et al.* 1993c; Santolini *et al.* 1994). Ces protéines sont impliquées dans la liaison aux récepteurs et corécepteurs dans la fusion membranaire, permettant l'entrée du virus dans la cellule. Ces protéines sont les cibles privilégiées de la recherche pour un vaccin anti VHC. Cependant ces deux molécules subissent un fort taux de mutation, rendant plus complexe la synthèse d'un vaccin prophylactique. De plus, la protéine E2 possède deux régions hypervariables appelées HVR1 et HVR2 (Satoru Saito *et al.* 1995).

Concernant le polypeptide **p7**, il appartient à la famille des viroporines et serait impliqué dans la maturation des particules virales en augmentant la perméabilité membranaire. Cette protéine de 63 acides aminés est localisée entre les protéines E2 et NS2 (Lin *et al.* 1994, Voisset & Dubuisson, 2004).

NS2 est une protéine de 180 aa, partiellement transmembranaire dont le rôle consiste au clivage protéolytique de la liaison entre NS2 et NS3.

NS3 possède deux activités enzymatiques participant à la fois à la réplication (fonction héli-

case nucléoside triphosphatase) ainsi qu'une fonction sérine protéase en présence de Mg^{2+} nécessitant le recrutement de NS4A comme co-facteur.

NS4A est un cofacteur de l'activité NS3 qui est également essentiel pour le clivage NS3/NS4A, NS4B/NS5A et également pour le clivage NS4A/4B et 5A/5B. Dirige et stabilise les protéines vers les membranes, sites de protéolyse.

NS4B possède actuellement une fonction encore hypothétique mais serait impliquée dans la formation du complexe de réplication (réseau membranaire) en induisant une altération membranaire (Gosert *et al* 2003).

NS5A est une protéine de 56 kDa dont la fonction reste encore à déterminer. Elle serait probablement impliquée dans la modulation de la réponse à l'interféron et interagirait avec de nombreux facteurs cellulaires.

NS5B est l'enzyme clef de la réplication du VHC, cette protéine possède en effet une activité ARN-polymérase-ARN-dépendante, qui assure la synthèse des deux brins d'ARN complémentaire.

Ces protéines sont la cible de nombreuses recherches pour une éventuelle thérapie antivirale spécifique du VHC (Yang *et al.* 2007).

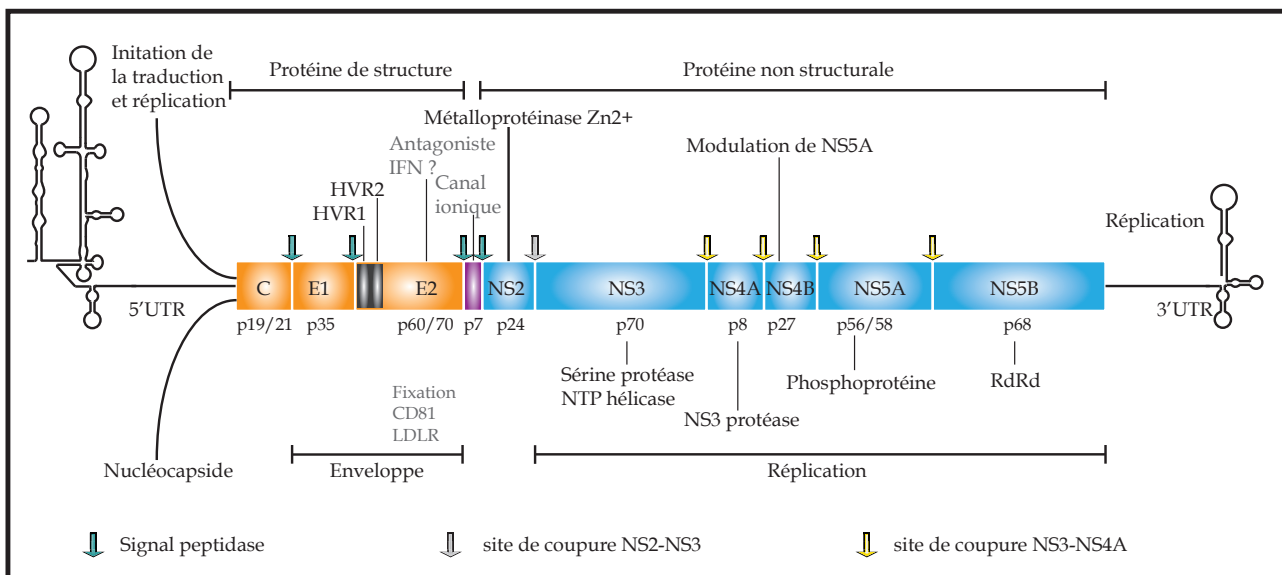


FIG7: Présentation détaillée des constituants de la polyprotéine du VHC.

III-4 Cycle de vie du VHC

Certaines parties du cycle de vie du virus demeurent encore mal connues, étant donné que, jusqu'à très récemment, aucun modèle efficace n'était encore disponible pour la réplication du virus. Deux récepteurs cellulaires ont été identifiés comme interagissant avec la protéine d'enveloppe E2 du VHC : la tétraspanine CD81 et le récepteur scavenger SR-B1. Après l'attachement du virus, celui-ci entre dans la cellule par endocytose avant la décapsidation du virus et la libération du génome viral (ARN positif simple brin) dans le cytoplasme. Le génome est directement traduit en une polyprotéine précurseur qui est maturée rapidement pour former les protéines virales actives. Parmi celles-ci, la réplicase NS3-5B synthétise l'ARN négatif qui, à son tour, sert de matrice pour la synthèse de grande quantités d'ARN viral positif. En fin de cycle, l'ARN viral (+) est encapsidé par le biais de l'interaction avec les protéines structurales et permet l'assemblage du virion. Les particules sont enveloppées en bourgeonnant dans la lumière du RE et sécrétées à l'extérieur de la cellule par le Golgi (Bartenschlager, Ahlborn-Laake *et al.* 1993; Lohmann, Korner *et al.* 1999; Watanabe *et al.* 2003; Bartenschlager *et al.* 2004, Dubuisson *et al.* 2008) (**FIG8**).

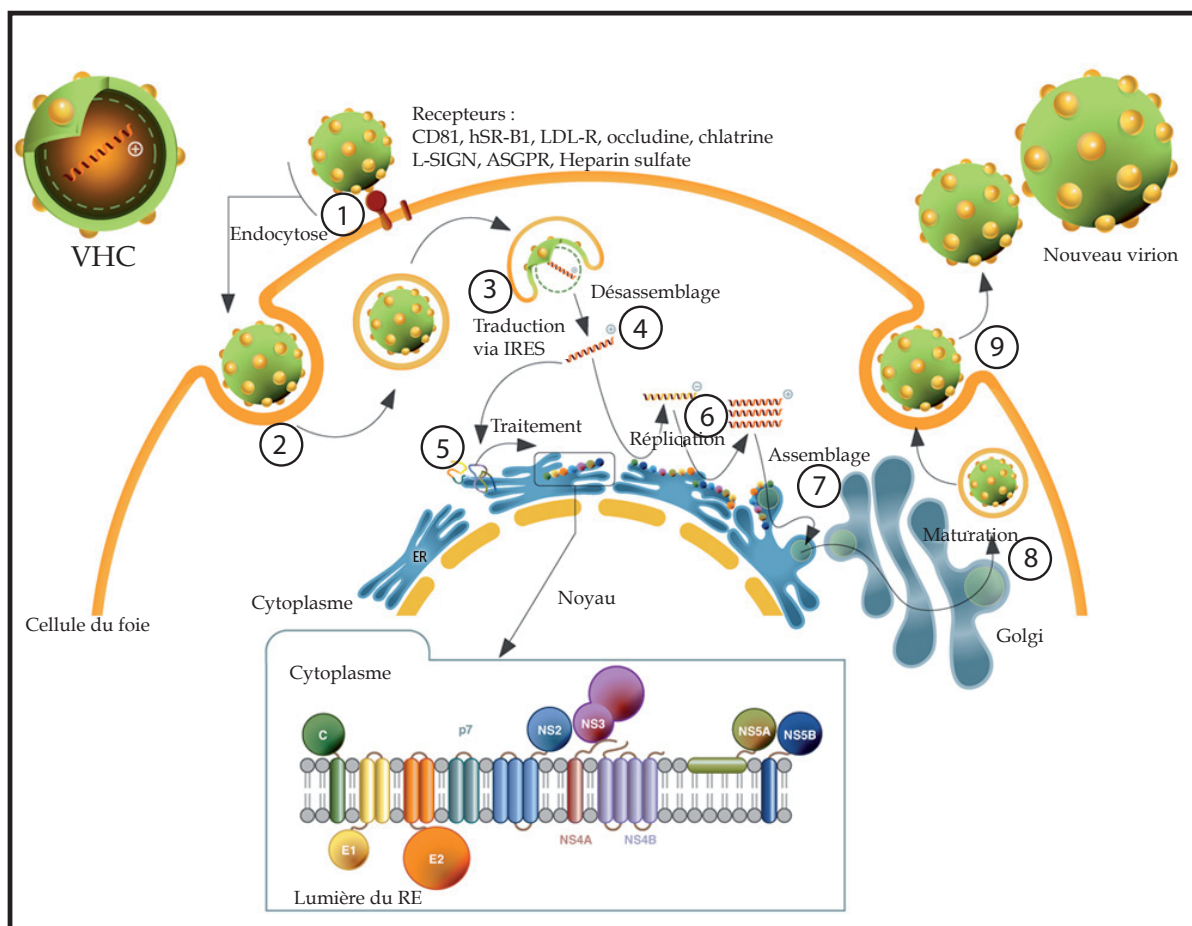


FIG8 : Cycle de vie du VHC.

- (1) Attachement
- (2) Entrée
- (3) Initiation de la traduction
- (4) Traduction
- (5) Maturation de la protéine
- (6) Réplication de l'ARN
- (7) Assemblage
- (8) Maturation
- (9) Nouveau virion

Source : http://www.tibotec.com/content/backgrounders/www.tibotec.com/hcv_lifecycle.html

III-5 Modèles pour l'étude du VHC

L'obstacle majeur pour l'étude du VHC est l'absence d'un système cellulaire fiable et efficace, c'est à dire capable d'induire une réplication complète du virus. Différentes approches ont été abordées par plusieurs équipes :

Modèles animaux : Un modèle *in vivo* théorique pour l'infection au VHC devrait permettre l'étude du VHC humain authentique, de l'interaction du virus avec les hépatocytes humains et de l'effet du système immunitaire de l'hôte sur les cellules infectées au VHC. Cependant, la plupart des modèles disponibles ne répondent que partiellement à ces critères. Le modèle animal le plus proche de l'homme est certainement le chimpanzé et plus généralement les primates. Cet hôte naturel développe toutefois une infection chronique à une fréquence moins élevée (< 50%) que les sujets humains (< 80%). Les souris transgéniques, quant à elles, ont permis l'étude de la réplication de l'expression génique et de l'immunopathogenèse du VHC. Mais il existe une limitation importante de ce modèle, ces souris sont en effet tolérantes aux antigènes viraux, contrairement aux sujets humains qui sont naïfs au VHC au moment de l'infection. Ces souris ne génèrent donc pas de réponse immune envers le virus et seule la réponse innée ainsi que la réponse des cytokines peuvent être étudiées; (revue : Guha *et al.* 2005).

Modèles cellulaires : En raison des problèmes évoqués ci-dessus, différents modèles cellulaires ont été développés afin de tenter d'étudier le VHC (Moriishi et Matsuura 2003; Guha *et al.* 2005). Pour des raisons encore incomprises aujourd'hui, l'infection d'hépatocytes primaires et de lignées cellulaires avec les différents génotypes viraux n'entraîne qu'une très faible réplication virale (Bartenschlager et Lohmann 2000). Le développement de systèmes cellulaires comprenant un réplicon subgénomique du VHC et supportant la réplication de l'ARN viral a permis l'élucidation des exigences pour la réplication du virus (Bartenschlager et Lohmann 2000; Blight *et al.* 2003). Cependant, il n'y a toujours pas de production de capsides du VHC (Blight *et al.* 2003). Récemment, un nouveau système de réplicon a été développé qui utilise une construction d'ADN du VHC entre deux ribozymes, conçus pour générer les extrémités 5'et 3' du VHC. Le réplicon subgénomique du VHC génotype 2a, JFH-1, a permis la culture *in vitro* et *in vivo* du VHC, permettant d'atteindre en culture des titres viraux de 10^5 unités infectieuses/ml en 48 heures (Kato *et al.* 2003; Lindenbach *et al.* 2005; Lindenbach *et al.* 2006).

Modèles *in vitro* : Différents modèles *in vitro* ont également été développés pour tenter de mieux comprendre les processus menant à l'assemblage viral et à l'encapsidation du génome du VHC. Par exemple, l'équipe de Klein a développé un système acellulaire, contenant un lysat de réticulocytes de lapin ou des extraits de germe de blé, exprimant la protéine de la capsid du VHC (Klein *et al.* 2004).

L'historique de la découverte du virus de l'hépatite C ainsi que sa biologie ayant été présentées, nous allons aborder la partie suivante dédiée à une étape clef qu'emprunte le virus pour traduire ses protéines alors que la traduction des protéines cellulaires est elle même bloquée.

Après un bref rappel du mécanisme classique de l'initiation de la traduction eucaryotes, j'aborderai une présentation détaillée du mécanisme particulier utilisé par le virus pour initier la traduction de ses protéines virales par l'intermédiaire d'une séquence d'ARN spécifique.

IV Initiation de la traduction dépendante de la coiffe

IV-1 Expression génique et traduction chez les Eucaryotes

La vie de la cellule eucaryote est régie par un ensemble de mécanismes complexes dépendant de l'expression et de la régulation de gènes spécifiques. L'ADN, contenant l'information génétique, est transcrit dans le noyau en ARN pré-messager par l'intermédiaire de nombreux facteurs. Ces ARN subissent, dès leur synthèse, des modifications post-transcriptionnelles, telles que l'épissage, la poly-adénylation ou l'ajout d'une coiffe (guanosine méthylée en position N7, reliée au premier nucléotide transcrit par une liaison 5'-5' triphosphate). Toutes ces étapes conduisent à l'obtention d'un ARN messager (ARNm) mature. La traduction étant effectuée dans le cytoplasme, les ARNm doivent donc changer de compartiment cellulaire par transport actif via les pores nucléaires. Les ARN messagers sortent sous la forme de particules : les hnRNPs (« heterogeneous nuclear ribonucleoparticles »). La traduction est orchestrée par de nombreux facteurs cellulaires et peut être décomposée en trois étapes : (i) l'initiation qui place le ribosome à l'aide des différents facteurs sur l'ARNm (ii) l'élongation qui consiste en la polymérisation des acides aminés reliés par une liaison peptidique et (iii) la terminaison, qui entraîne le relargage du peptide néo-synthétisé dans le cytoplasme.

IV-2 L'initiation et la reconnaissance de la coiffe de l'ARNm

La reconnaissance de la coiffe implique la présence de plusieurs facteurs eucaryotes qui conduisent au recrutement de la sous-unité 40S du ribosome (**FIG9**).

IV-2-A Formation du complexe ternaire eIF2•GTP•ARNtiMet

L'initiation de la traduction commence par la formation du complexe ternaire constitué par le facteur eIF2, une molécule de GTP et un ARN de transfert initiateur chargé de méthionine (ARNtiMet) : eIF2•GTP•ARNtiMet (Kapp et Lorsch 2004).

IV-2-B Formation du complexe de pré-initiation 43S

Après cette première étape, la sous-unité ribosomique 40S, en présence de facteurs d'initiation va se lier au complexe ternaire (Kapp et Lorsch 2004). Les facteurs eIF1, eIF1A, eIF3 et eIF5 sont les facteurs nécessaires à ce recrutement (Majumdar *et al.* 1979; Chaudhuri *et al.* 1997; Majumdar *et al.* 2003). Le facteur eIF1A joue un rôle dans l'association de l'ARNtiMet à la sous-unité ribosomique 40S. Les facteurs eIF3 et eIF1A quant à eux, empêchent les deux

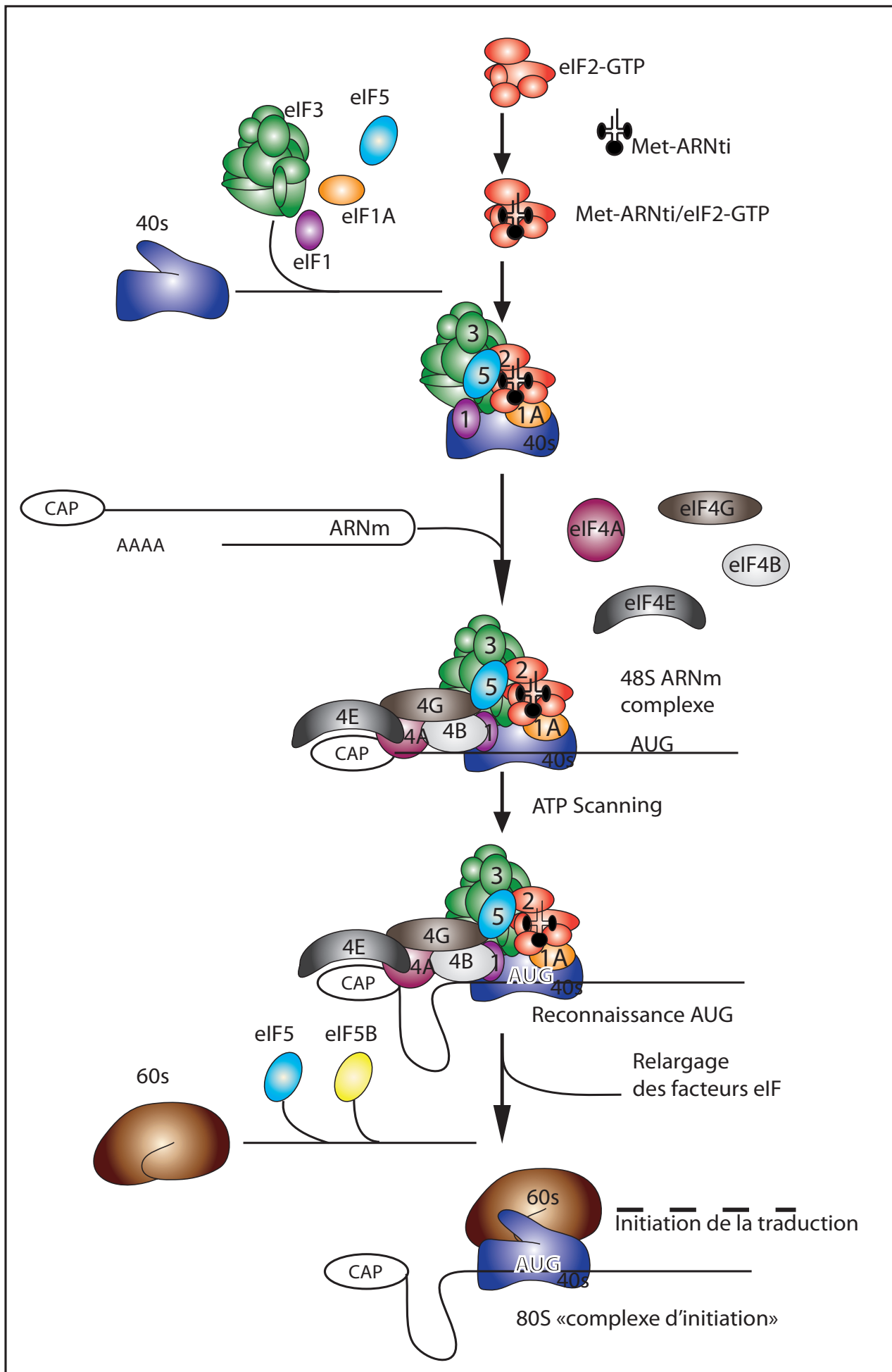


FIG9 : Organisation de l'initiation de la traduction coiffe dépendante.

sous-unités ribosomiques de s'associer de façon prématurée, ce qui conduirait à un ribosome non fonctionnel (Hershet et Merrick, 2000). La présentation du facteur eIF3 sera abordée en détail au paragraphe suivant.

IV-2-C La coiffe est reconnue par le complexe eIF4F

Le facteur multiprotéique eIF4F, composé de 3 protéines distinctes, eIF4E, eIF4A et eIF4G pilote l'attachement du complexe 43S sur l'ARNm. eIF4E lie spécifiquement la coiffe méthylée en 5' de l'ARNm. eIF4G assure la cohésion de l'ensemble car il interagit avec eIF4E (côté ARN) et avec eIF3 (côté ribosome). eIF4A, quant à lui, déroule les structures secondaires de l'ARN grâce à son activité hélicase. Outre son interaction avec eIF4E, eIF4G peut se lier à eIF4A, eIF3 et PABP (« poly-A binding protein ») eIF4F est une protéine d'échafaudage qui permet de rassembler des facteurs capables d'interagir à la fois avec la coiffe mais aussi avec la queue poly(A) de l'ARNm afin d'augmenter de façon synergique l'efficacité de l'initiation (Prevot, Darlix *et al.* 2003). Ce complexe est la cible de certaines protéases virales qui, en dégradant ce complexe, inactive très efficacement la synthèse *de novo* de protéines.

IV-2-D Le balayage

Une fois le complexe 48S formé, l'ARNm va être balayé jusqu'à la rencontre d'un codon initiateur. Ce mécanisme nécessite de l'énergie en utilisant l'hydrolyse de l'ATP. La variation, dans le choix du codon d'initiation au sein d'un même ARNm, contribue à la diversité des protéines générées (Touriol *et al.* 2003). Dans 90% des cas, l'initiation a lieu au niveau d'un codon d'initiation AUG (Kozak 1980). Des structures secondaires de l'ARN en épingles à cheveux situées en aval d'un codon AUG augmentent l'efficacité de l'initiation au niveau de ce codon (Kozak 1990). Dans le cas où le contexte n'est pas favorable, le ribosome passe parfois sans initier la traduction jusqu'à ce qu'il rencontre un autre codon situé dans un contexte plus favorable, c'est le « leaky scanning » (Kozak 1997).

IV-2-E Assemblage du ribosome 80S et fin de l'initiation

La formation du complexe final 80S correspond à la phase où les autres facteurs d'initiation sont expulsés, permettant ainsi à la grande sous-unité ribosomique 60S de se lier à la petite sous-unité 40S (Fraser et Doudna 2007). Cette étape finale nécessite la présence du facteur eIF5B•GTP pour catalyser cette liaison (Pestova *et Hellen* 2000; Pestova *et al.* 2000; Kolupaeva *et al.* 2005). Le facteur eIF6 qui se lie à la sous-unité 60S du ribosome joue un rôle important en empêchant, à l'image des facteurs eIF1 et eIF3, l'association des deux sous-unités ribosomiques de façon prématurée. L'ARN^{ti}Met peut donc occuper le site P du ribosome et l'élongation de la chaîne polypeptidique débute. Cette dernière étape représente la fin de la phase d'initiation.

IV-3 L'élongation

Dès la mise en place du ribosome, l'élongation, qui correspond à l'incorporation séquentielle des acides aminés, commence (revue Merrick *et Nyborg*, 2000). C'est un phénomène cyclique régulé par la présence des facteurs protéiques eEF (eukaryotic Elongation Factors). Le facteur eIF2, couplé au GTP, interagit avec l'ARNtⁱ Met, et perd son affinité pour la méthionine après hydrolyse du GTP ce qui libère l'ARNtⁱ. Le facteur eIF2 quitte alors le ribosome sous la forme d'un complexe binaire eIF2-GDP. Le facteur eIF2 peut-être recyclé par action du facteur eIF2b qui catalyse l'échange du GDP en GTP (Kapp *et Lorsch* 2004). Ce phénomène permet d'incorporer entre 6 à 8 acides aminés par seconde, soit environ 1 minute pour une protéine de 50 kDa. Le ribosome possède trois sites de liaison à l'ARNt :

Le site A (AminoacylARNt) où arrive l'ARNt chargé de l'acide aminé correspondant.

Le site P (PeptidoacylARNt) où l'acide aminé est lié par liaison peptidique à la chaîne peptidique en cours.

Le site E (Exit) permet l'expulsion du ribosome de l'ARNt-déacétylé.

Après translocation, l'ARNt passe du site A au site P puis au site E, le site A devenu inoccupé peut accepter une nouvelle molécule d'ARNt-acide aminé, et le cycle recommence.

IV-4 La terminaison

La fin de l'élongation correspond à l'arrivée d'un des trois codons STOP (UAA, UAG ou UGA) au niveau du site A. La présence de ce codon au ribosome induit la reconnaissance des facteurs de terminaison composé de facteurs eRF (eukaryotic Releasing Factors). Une fois la chaîne polypeptidique libérée, les facteurs quittent le ribosome qui se sépare de l'ARNm par dissociation des deux sous-unités.

V Initiation de la traduction Coiffe indépendante

L'initiation de la traduction selon le modèle de balayage coiffe-dépendant du ribosome proposé par Marilyn Kozak en 1980 permet d'expliquer le mécanisme traductionnel de la majorité des ARN messagers (ARNm) cellulaires eucaryotes. Il est important de noter que 3% des ARNm cellulaires utilisent une voie indépendante de la présence d'une coiffe et qui sont généralement impliqués dans des processus de survie cellulaire. L'étude d'un tel mécanisme a tout d'abord été identifié chez les Picornavirus ce qui a permis de mettre en évidence la présence d'une séquence (appelée séquence IRES pour Internal Ribosome Entry Site) qui permet à l'ARN non coiffée d'être traduit. Ce type d'ARN a depuis été découvert chez de nombreux virus ainsi que dans certains ARNm cellulaires. Les IRES permettent aux virus de détourner la machinerie classique de traduction cellulaire à son profit.

Les ARNm des *Picornaviridae* (virus à ARN de polarité positive) possèdent une longue région 5'UTR très structurée et dépourvue d'une coiffe à leur extrémité. Cette région contient des

codons d'initiation AUG situés en amont du codon d'initiation canonique, mais qui ne sont pas utilisés pour la traduction. Toutes ces caractéristiques vont à l'encontre du modèle classique d'initiation de la traduction par balayage coiffe-dépendant. L'infection par la plupart des *Picornavirus* conduit à un arrêt rapide (5h après l'infection) de la traduction des ARNm cellulaires, suite au clivage en deux parties du facteur d'initiation eIF4G par des protéases virales (2A ou L) (Ohlmann *et al.* 1997). Le fragment amino-terminal contient le site de liaison à eIF4E, tandis que le fragment carboxy-terminal contient les sites de liaison à eIF3 et eIF4A. Ainsi, la séparation entre la fonction de reconnaissance de la coiffe (eIF4E) et les domaines d'interaction avec l'hélicase eIF4A et le facteur eIF3 conduit à l'inhibition de la traduction coiffe-dépendante des ARNm cellulaire. La présence du virus induit une hypophosphorylation du facteur de 4E-BP qui reste lié à eIF4E et renforce l'inhibition de la synthèse des protéines cellulaire. En revanche, la traduction de l'ARN viral n'est pas inhibée car elle est IRES-dépendante, et peut utiliser le fragment C-terminal de eIF4G qui est associé à eIF4A, et utiliser le facteur eIF3 et le ribosome. Le modèle simplifié de l'initiation de la traduction coiffe indépendante orchestré par l'IRES du VHC est présenté dans la **figure 10**.

(1) formation du complexe ribosomique 40S + IRES.

L'IRES est capable de recruter directement la sous-unité ribosomique 40S sans la présence d'autres acteurs protéiques. Une fois cette étape effectuée, les facteurs eIF3 et eIF2-GTP sont mobilisés pour former le pré-complexe d'initiation 43S.

(2) formation du complexe de pré-initiation (IRES/40S + eIF3 et Met-ARNt-eIF2-GTP).

Après une phase impliquant l'hydrolyse du GTP, les facteurs eIFs sont alors remplacés par la sous unité 60S pour former le complexe 80S (Pestova *et al.* 1998).

(3) formation du complexe d'initiation 80S (IRES/40S avec élimination des eIFs + sous-unité ribosomale 60S) (Rijnbret *et al.* 1996; Fraser *et al.* 2004; Otto et Puglisi 2004).

Plusieurs modèles de l'initiation de la traduction coiffe indépendante sont disponibles dans la littérature, et rapportent le même type de déroulement. (Pour une revue cf. Fraser et Doudna, 2007 Nature).

La comparaison des mécanismes d'initiation coiffe dépendante / coiffe indépendante du VHC nous indique la présence de différences essentiellement basées sur le nombre de facteurs cellulaires recrutés. On peut remarquer que le nombre de facteurs requis dans l'initiation de la traduction coiffe dépendante est nettement supérieur à celui utilisé dans le cas de l'initiation indépendant de la coiffe. En effet, toutes les protéines reconnaissant la coiffe ainsi que celles assurant la liaison avec eIF3 ne sont pas impliquées dans la traduction coiffe indépendante (**Tableau III**).

Tableau III : Récapitulatif des facteurs nécessaires à l'initiation de la traduction

Nom	Nom des facteurs
Points communs	40S, eIF3, eIF2, 60S
Coiffe dépendant	eIF1/1A, eIF2, eIF2B, eIF4A/B/G/E, eIF5A/B

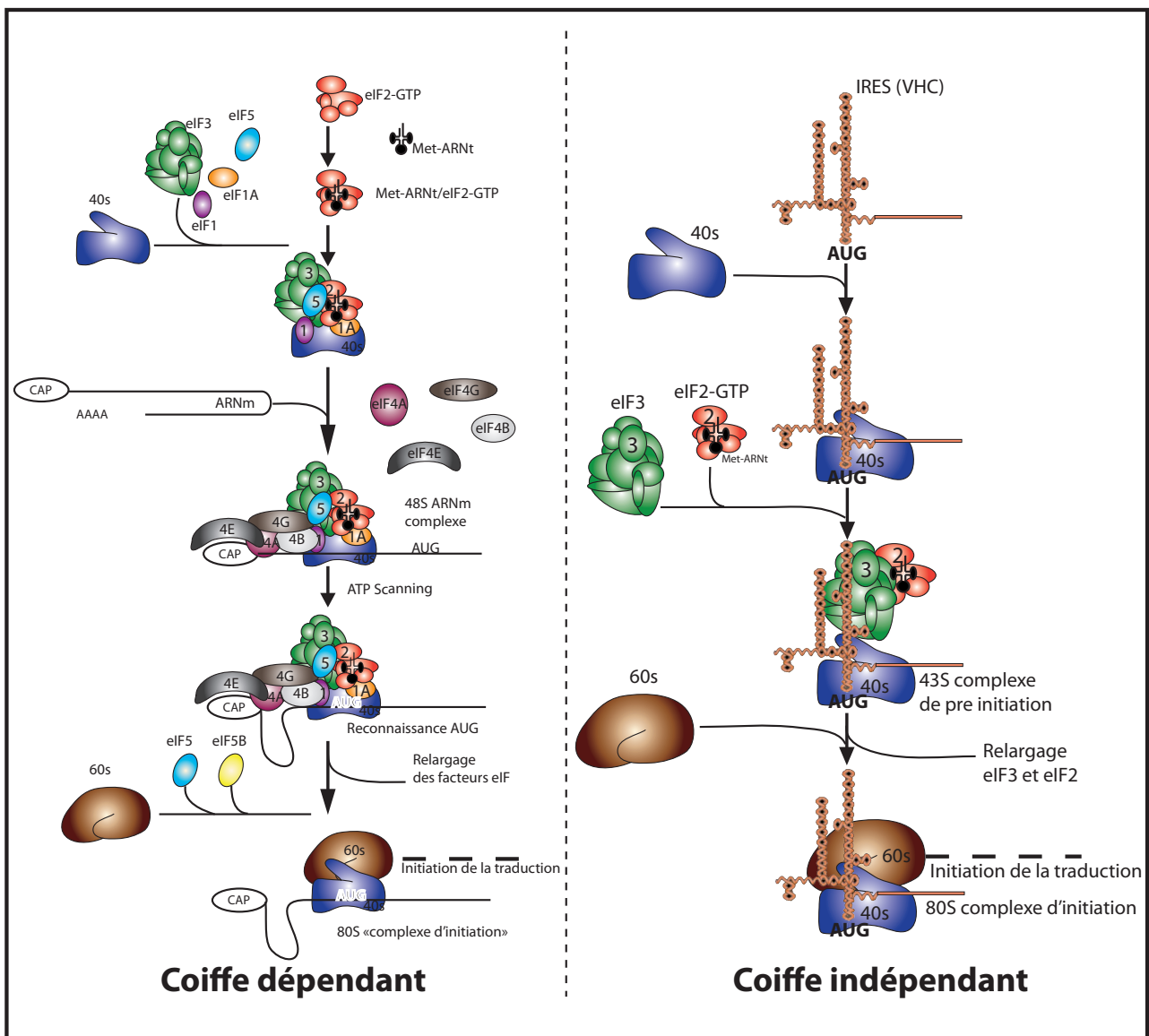


FIG10 : Déroulement de l'initiation de la traduction coiffe indépendante (cas du VHC) en comparaison de la traduction coiffe dépendante.

VI Un facteur commun : eIF3

Le facteur eIF3 a été identifié dans les années 1970 (Issinger *et al.* 1976; Majumdar *et al.* 1979) en tant que stimulateur de la synthèse protéique, le facteur de démarrage eIF3 commence à être de mieux en mieux compris, quant à son organisation et ses principales fonctions. Ce facteur est l'élément le plus complexe de la machinerie traductionnelle eucaryote avec une masse de 750kDa chez l'humain. Ce facteur comporte 13 sous-unités chez les mammifères appelées eIF3a à eIF3m (Asano *et al.* 1997; Browning *et al.* 2001; Mayeur *et al.* 2003; Damoc *et al.* 2007). Chez la levure, le complexe eIF3 ne comporte que 5 sous-unités dont les noms ont été unifiés avec ceux utilisés pour le facteur humain : TIF32/3a, PRT1/3b, NIP1/3c, TIF34/3i et TIF35/3g (TIF : Translation Initiation Factor, PRT : PRoTein synthesis, NIP : Nuclear ImPort). Le facteur eIF3 possède de nombreuses fonctions dans l'initiation de la traduction et induit l'assemblage du complexe ternaire ARNtMet-eIF2-GTP. Il est aussi capable d'interagir directement avec la sous-unité ribosomique 40S en absence d'autres facteurs (Pestova et Hellen 2001). De plus, ce facteur induit la formation du complexe de pré-initiation 43S et permet de maintenir dans un état dissocié les sous-unités 40S et 60S avant que le complexe 48S ne soit complètement formé (Kolupaeva *et al.* 2000; Kolupaeva *et al.* 2005; Kolupaeva *et al.* 2007). Au cours du démarrage de l'initiation de la traduction coiffe dépendante, le complexe eIF3 organise un vaste réseau d'interactions avec de nombreux partenaires parmi lesquels : eIF1, eIF1A, eIF2, eIF5, eIF4G, eIF4B et la sous-unité 40S (Methot *et al.* 1996; Methot *et al.* 1997; Bandyopadhyay et Maitra 1999; Fletcher *et al.* 1999); Pour une revue : (Hinnebusch 2006).

Le facteur eIF3 dans le cadre de l'initiation indépendante de la coiffe est impliqué dans un réseau moins étendu et rassemble eIF2, 40S et l'IRES. De plus, on pourra retrouver dans certains

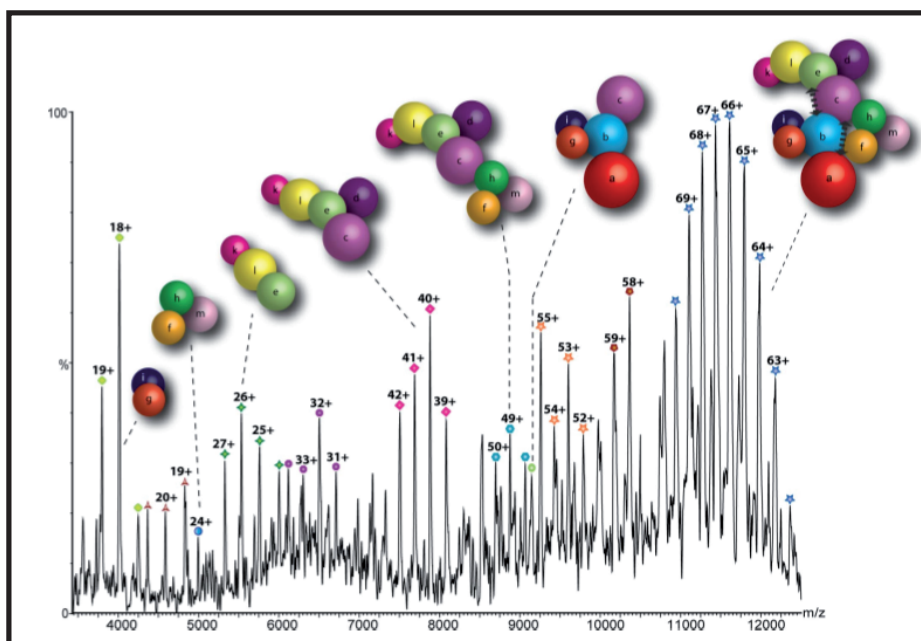


FIG11 : Analyse des sous-unités du complexe eIF3 humain.

Etude des interactions intramoléculaires des sous-unités du complexe eIF3 par spectrométrie de masse (Zhou *et al.* 2008).

cas des protéines ITAF dont le rôle reste encore à définir.

Des travaux, visant à identifier le noyau central du complexe eIF3 eucaryote, ont été réalisés par l'équipe de Masutani (Masutani *et al.* 2007) et ont permis d'identifier six sous-unités indispensables au fonctionnement du complexe (eIF3a, eIF3b, eIF3c, eIF3e, eIF3f, et eIF3h). Ce cœur est suffisant pour réaliser une interaction avec la sous-unité 40S en présence d'ARNm, mais toutes ces sous-unités n'interagissent pas avec l'ARN ribosomique (cf tableau IV). Une autre approche a été abordée par l'équipe de Zhou (Zhou *et al.* 2008) qui a effectué un travail important sur l'identification des interactions protéine/protéine reliant les différentes sous-unités au sein du complexe (**FIG11**). On peut, à la suite de ces travaux, mieux comprendre l'arrangement moléculaire des différentes sous-unités. On remarque que seule la sous-unité b interagit avec cinq des douze sous-unités, faisant de cette protéine la plate-forme d'échafaudage du complexe (**Tableau IV**).

Tableau IV : Récapitulatif des sous-unités du complexe eIF3 et de leurs interactions

Nom	Masse kDa	Interaction au sein du complexe eIF3	Interaction avec de l'ARN ribosomique	Interaction avec un IRES
eIF3a	170	b, c	oui	oui
eIF3b	116	g, i, c, a, e	oui	oui
eIF3c	110	e, h	oui	oui
eIF3l	69	k, e	n.a	n.a
eIF3d	66	e	oui	oui
eIF3e	48	l, d, c	n.a	n.a
eIF3f	47	h, m	oui	oui
eIF3g	44	i, b	oui	n.a
eIF3m	42	h	n.a	n.a
eIF3h	40	c, f, m	n.a	n.a
eIF3i	36	g, b	n.a	n.a
eIF3j	35	b	oui	n.a
eIF3k	28	l, d, c	n.a	n.a

Les données structurales à haute résolution de ce facteur multiprotéique ne représentent que 10% du complexe : sous-unités eIF3k, eIF3j, MRR-eIF3b et MRR-eIF3g. Les données structurales sont présentées dans la **figure 12**.

Les domaines MRR-eIF3g (aa 239-317, pdb : 2CQ0) et MRR-eIF3b (aa 185-268, pdb : 2NLW) résolu par la technique de RMN, sont composés de deux hélices α et de cinq feuilletts β , suivant un agencement $\beta\alpha\beta\beta\alpha\beta$ caractéristique d'un Motif de Reconnaissance de l'ARN (MRR). On peut remarquer que ces deux motifs possèdent une forte homologie de structure (95%) mais ne possèdent pas les mêmes fonctions. En effet, le motif MRR porté par la sous-unité g a été caractérisé comme impliqué dans une interaction avec l'ARN ribosomal de la sous-unité 40S

tandis que le motif MRR de la sous unité b a été décrit comme impliqué dans des interactions avec des protéines.

La protéine eIF3k (aa 1-226) est spécifique des eucaryotes supérieurs et possède un rôle régulateur avec une interaction avec les sous-unités eIF3c, eIF3g et eIF3j (Mayeur *et al.* 2003).

De récents travaux impliquent la sous-unité eIF3j (Fraser *et al.* 2004; Fraser *et al.* 2009) dans la formation du complexe eIF3/40S ainsi que dans l'interaction eIF3/ARNm (Fraser *et al.* 2007; Fraser et Doudna 2007) dans un cadre dépendant de la présence d'une coiffe. L'étude des différentes sous-unités semble être complexe et de nombreuses équipes ont eu recours à la microscopie électronique afin d'obtenir des informations sur ce facteur. Grâce à cette technique, il a été possible d'apprécier la structure polylobée de ce complexe, qui semble posséder une structure allongée de 250 à 300Å. Cependant, à cette résolution (20Å), il n'est pas possible d'identifier ou de localiser les sous-unités le composant. On peut remarquer que la plupart des informations structurales disponibles présentent le facteur eIF3 dans le cadre de l'initiation de la traduction dépendante de la coiffe. Seule une étude en microscopie électronique étudiant le complexe eIF3/IRES du VHC a permis de montrer la forme du facteur eIF3 au sein du complexe. Ces résultats indiquent que le facteur eIF3 possède la même forme en complexe avec l'IRES qu'en complexe avec la sous-unité 40S. L'IRES n'influerait donc pas sur la structure du facteur eIF3.

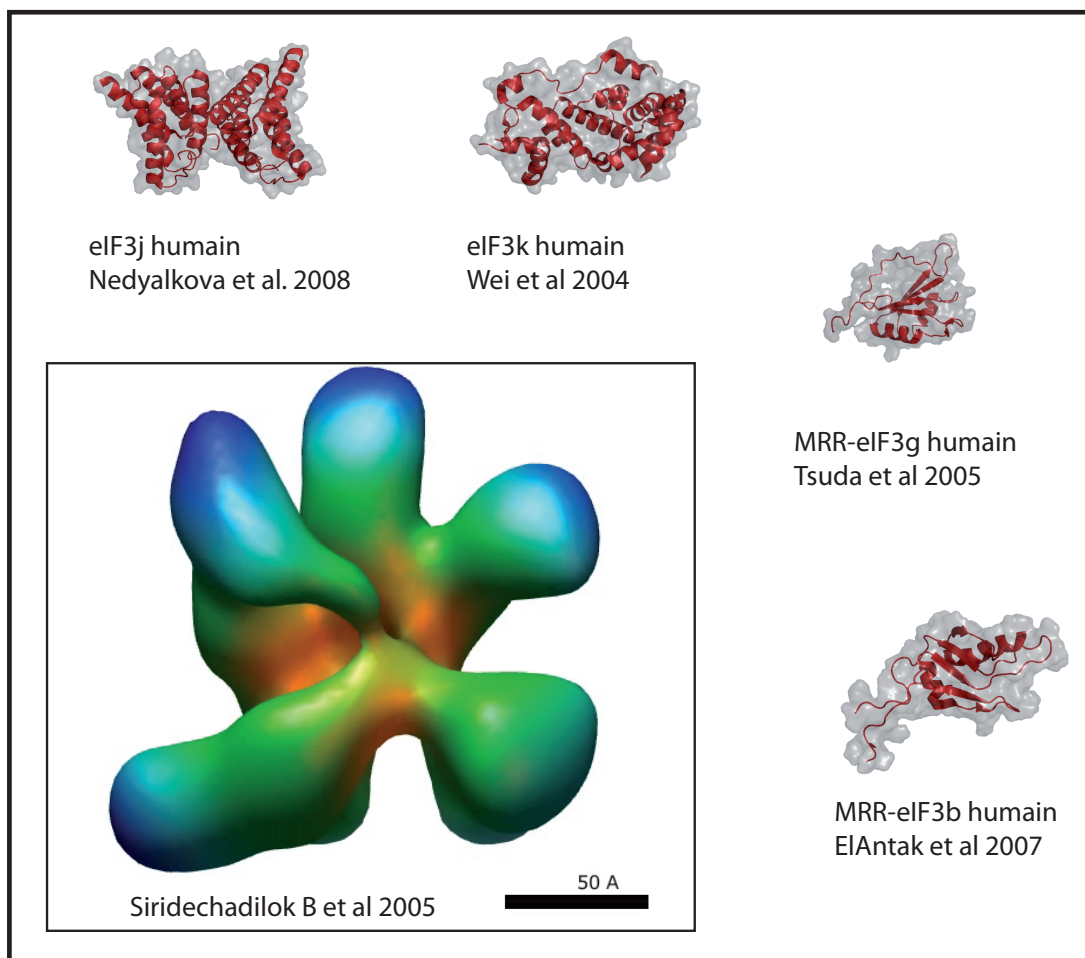


FIG12 : Présentation des données structurales du complexe eIF3.

VII Les IRES

Les premiers IRES ont été découverts, il y a plus de dix ans chez les *Picornavirus*. Depuis, ils ont été décrits chez de nombreux autres virus à génome ARN de polarité positive, certains rétrovirus et même chez des virus à ADN (Mokrejs *et al.* 2006) (**Tableau V**). Plus récemment, des séquences IRES ont été également décrites chez des ARNm cellulaires. La plupart de ces ARNm cellulaires codent des protéines indispensables à la survie cellulaire en cas de stress, lorsque la traduction par voie classique est compromise (pour une revue : Stephen D.Baird 2006).

La description générale des IRES semble complexe et il ne semble pas aisé de se baser uniquement sur des analyses de séquences pour les caractériser. C'est un regard global sur ces différents mécanismes qui nous permettra d'apprécier toute la complexité, mais aussi toute la simplicité de ce système.

Tableau V : Liste non exhaustive des virus possédant un IRES

Nom		Référence
Virus à ARN simple brin positif		
<i>Picornaviridae</i>	Enterovirus : PV	(Pelletier <i>et al.</i> 1988)
	Rhinovirus : RVH	(Borman <i>et al.</i> 1992)
	Hepatovirus : VHA	(Glass <i>et al.</i> 1993)
	Cardiovirus : EMCV	(Jang <i>et al.</i> 1988)
	TMEV	(Bandyopadhyay <i>et al.</i> 1992)
	Aphrovirus : FMDV	(Belsham <i>et al.</i> 1990)
<i>Dicistroviridae</i>	Dicistrovirus : CrPV	(Wilson <i>et al.</i> 2000)
	TSV	(Jang, Lo <i>et al.</i> 2009)
<i>Flaviviridae</i>	Pestivirus : BVDV	(Poole, Wang <i>et al.</i> 1995)
	CSFV	(Rijnbret <i>et al.</i> 1997)
	Hepacivirus : VHC	(Tsukiyama-Kohara, Iizuka <i>et al.</i> 1992)
<i>Potyviridae</i>	Potyvirus : TEV	(Gallie <i>et al.</i> 2001)
Virus à transcriptase inverse		
<i>Retroviridae</i>	Lentivirus : VIS	(Ohlmann <i>et al.</i> 2000)
	VIH	(Buck <i>et al.</i> 2001)
	Retrovirus HTLV : VTLH	(Attal <i>et al.</i> 1996)
	Retrovirus type C : VLSM	(Vagner, Gensac <i>et al.</i> 1995)
	VSR	(Deffaud <i>et al.</i> 2000)
Virus à ADN double brin		
<i>Herpesviridae</i>	Gammaherpesviridae : VSK	(Grundhoff <i>et al.</i> 2001)

VII-1 Les IRES cellulaires

La découverte du premier IRES cellulaire fait suite à l'étude de la protéine BiP (Johannes et Sarnow 1998); Aujourd'hui 85 IRES ont été caractérisés sur des ARNm cellulaires (Baird, Turcotte *et al.* 2006), que l'on retrouve dans plusieurs voies biologiques (**Tableau VI**). On retrouve des séquences IRES sur les ARNm du facteur d'initiation de la traduction DAP5, ou eIF4G (Johannes et Sarnow 1998) sur des facteurs de croissance FGF2 (Vagner *et al.* 1995) ou encore potentiellement sur des protéines nucléaires comme l'Histone H4 (Gunjan *et al.* 2005). La présence de ces structures remet donc en question la régulation de la production de protéine au cours de la vie d'une cellule et particulièrement en cas de stress cellulaire. Dans le cas de la mitose par exemple, seulement 3% des ARNm cellulaires restent encore associés aux ribosomes, pourtant la présence d'une activité de traduction a été détectée pendant cette phase. De plus, il a été démontré qu'il existait une régulation de l'initiation indépendante de la coiffe des ARNm cellulaires pendant la mitose (Qin et Sarnow 2004). Comme évoqué plus haut, les initiations régulées par une séquence IRES ont généralement lieu dans des situations de stress cellulaire (Van Eden *et al.* 2004), par exemple dans l'angiogénèse (Stoneley et Willis 2004), la mort cellulaire en générale (Holcik *et al.* 2000), l'apoptose (Holcik et Sonenberg 2005) ou dans certains cancers (Holcik *et al.* 2004). Ces conditions de stress cellulaire conduisent souvent à un manque de facteurs canoniques. Lors de l'apoptose, l'utilisation d'une voie détournée impliquant les IRES permet la synthèse de facteurs impliqués dans la mort cellulaire programmée, de facteurs de survie et d'oncogènes (Holcik *et al.* 2000). Cependant ces ARNm réagissent dans les conditions normales comme des ARNm classiques. Un facteur déclenchant a été identifié permettant d'activer, d'une certaine manière, le système de secours de la traduction eucaryote.

Tableau VI : Liste non exhaustive des différents IRES cellulaires

	Protéine	Références
Expression génique	Facteurs de transcription	c-Myc (Nanbru, Lafon <i>et al.</i> 1997)
		<i>Antennapedia</i> (Droso) (Oh <i>et al.</i> 1992)
		<i>Ultrabithorax</i> (Ye <i>et al.</i> 1997)
	Maturation de l'ARN	La (Carter et Sarnow 2000)
	Facteurs de traduction	eIF4GI (Johannes et Sarnow 1998)
		TIF4631 (Zhou <i>et al.</i> , 2001)
P97/DAP5 (Henis-Korenblit <i>et al.</i> 2000)		
Réponse au stress	GRP78/BiP (Macejak et Sarnow 1991)	
Facteur de croissance	FGF2 (Vagner, Gensac <i>et al.</i> 1995)	
Cytosquelette	MAP2 (Pinkstaff <i>et al.</i> 2001)	
Kinase	P58 (Cornelis <i>et al.</i> 2000)	
Transport	Cat-1 (Fernandez <i>et al.</i> 2001)	
Assemblage des Histones	SLBP (Zhao <i>et al.</i> 2004)	

La traduction indépendante de la coiffe est en effet stimulée par le produit du clivage de l'homologue p97 du facteur eIF4G, p86 (Marash et Kimchi 2005). En revanche, la présence du facteur eIF4E semble diminuer l'efficacité de l'initiation indépendante de la coiffe (Hernandez *et al.* 2004). Les IRES cellulaires décrits jusqu'à présent ne possèdent pas de caractéristiques communes aisément identifiables avec les IRES viraux. On peut cependant établir que ces IRES cellulaires seraient beaucoup moins efficaces que les IRES viraux impliquant la présence d'un mécanisme encore incompris qui nécessiterait, d'après de récents travaux, une étape d'activation nucléaire (Gonzalez-Herrera *et al.* 2006). Les IRES cellulaires constituent une voie de secours pour la cellule conduisant à la synthèse de protéines impliquées dans la survie cellulaire. Une autre fonction leur est attribuée, celle d'intervenir dans la synthèse de plusieurs isoformes de protéines qui ont des activités biologiques différentes (Nanbru *et al.* 1997). Ce mécanisme a été identifié par l'équipe de Lloyd (Byrd *et al.* 2005) concernant l'ARNm de eIF4GI et permet donc la synthèse de cinq isoformes différents (notés de a à e) possédant des fonctions différentes, encore non élucidées, dans la régulation de la sous-unité eIF4G.

VII-2 Les IRES viraux

Pour mener à bien l'initiation de la traduction coiffe indépendante, l'intégralité de la séquence IRES est nécessaire. La taille des séquences est très variable selon les virus concernés. On peut remarquer que tous les IRES sont constitués de différents domaines repliés de manière indépendante, dont le nombre, la composition et la disposition varient.

VII-2-A Caractéristiques et diversités structurales

La définition d'une séquence IRES ne répond qu'à un seul critère fonctionnel : sa capacité à initier la traduction de manière indépendante de la coiffe. La taille des séquences IRES est très variable, allant en moyenne de 200 nucléotides pour les IRES des virus à ARN d'insectes, à environ 450 pour les IRES des *Picornavirus* (Martínez-Salas *et al.* 2001). Cependant, quelques caractéristiques communes entre ces séquences sont retrouvées, comme la grande quantité de structures secondaires et tertiaires qui semble essentielle pour le recrutement du ribosome. En effet, des délétions, des insertions ou même des substitutions d'un seul nucléotide affectant la structure secondaire ou tertiaire de l'IRES peuvent réduire et même abolir l'activité traductionnelle. La structure secondaire posséderait une fonction d'échafaudage capable de recruter la sous-unité ribosomale 40S par de nombreuses interactions intermoléculaires. Il est possible de classer les IRES viraux en trois familles.

VII-2-A-1 Première famille d'IRES :

L'analyse de la séquence des IRES de la famille des *Picornavirus* montre une séquence IRES de 450 nt qui débute à une centaine de nucléotides après l'extrémité 5'. Ils possèdent un

motif conservé : une courte région, riche en pyrimidines, située 25 nucléotides en amont de l'extrémité 3' de l'IRES (Pestova *et al.* 1991) (**FIG13**). Cette famille est divisée en trois sous-familles : deux groupes majeurs et un groupe mineur (Jackson *et al.* 1994). La division en sous-famille est basée sur la position de la séquence IRES par rapport au codon d'initiation situé soit en aval soit en amont de ce codon. Dans certains cas, on peut même retrouver plusieurs codons d'initiation qui sont indispensables à l'initiation sans être traduits. Les deux groupes majeurs partagent deux motifs en commun, une tétraboucle GNRA (où N est un nucléotide quelconque et R une purine), et une structure riche en cytosines (Fernandez-Miragall *et al.* 2009; Belsham *et al.* 1992; Belsham *et al.* 2009).

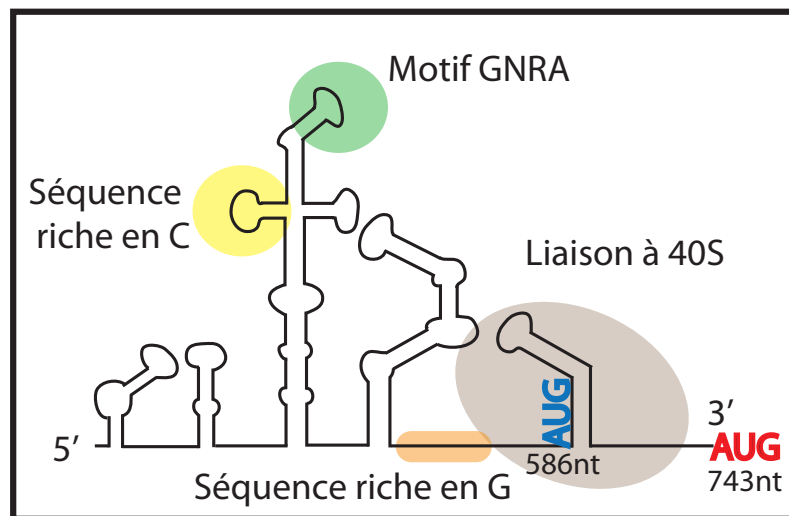


FIG13 : Représentation schématique des IRES de la famille des Picornavirus (D'après la prédiction de structure secondaire).

Illustration inspirée de : La synthèse protéique chez des virus animaux à ARN positif par un mécanisme alternatif d'initiation de la traduction (Malnou *et al.* 2002).

VII-2-A-2 Deuxième famille d'IRES :

Elle est représentée par les IRES des *Flaviviridae* (famille du Virus de l'Hépatite C). Les IRES de cette famille ont une longueur d'environ 350 nucléotides mais ne présentent qu'une similitude de séquence d'environ 50% (Le *et al.* 1998) (**FIG14**). Ils débutent à une vingtaine de nucléotides de l'extrémité 5' de l'ARN et s'étendent jusqu'à une trentaine de nucléotides après le codon AUG. Les nucléotides au sein de l'IRES sont fortement conservés par rapport à ceux situés dans le reste de l'ORF (Reynolds *et al.* 1995). Comme les autres IRES, des séquences riches en pyrimidines sont identifiables, mais ils présentent la particularité d'avoir l'AUG initiateur au niveau d'un pseudo-nœud essentiel au démarrage de la traduction.

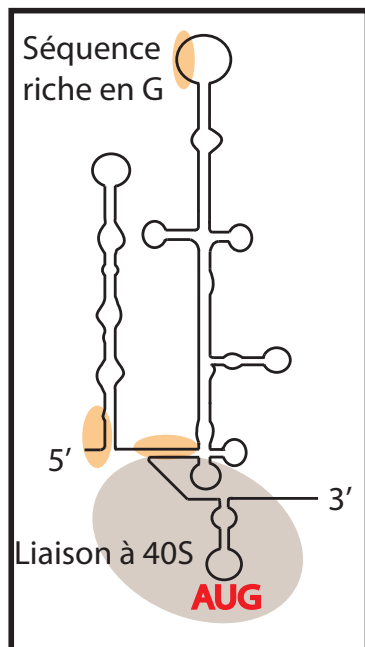


FIG14 : Représentation schématique des IRES de la famille des Flaviviridae.

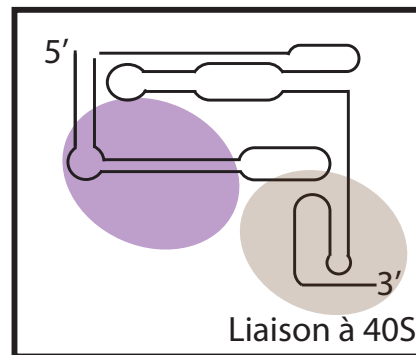


FIG15 : Représentation schématique de l'IRES de la famille des Dicistrovirus.

En violet : boucle L1.1 de la famille des Dicistrovirus
Illustration inspirée de : Simon P. Fletcher et Richard J. Jackson.

VII-2-A-3 Troisième famille d'IRES:

Elle est formée par les IRES des *Dicistrovirus*, auxquels appartient le Virus de la Paralysie du Cricket (CrPV). Ces IRES ont une longueur assez courte (280nt) et possèdent de nombreuses similarités avec les IRES des *Picornavirus*. Les deux principales caractéristiques de ces IRES résident dans le fait qu'ils présentent une structure analogue à un ARN de transfert (Kieft *et al.* 2008), et d'autre part, dans le fait que le codon initiateur AUG soit remplacé par un codon GCU (Pisarev *et al.* 2005) (**FIG15**).

De nombreuses études *in vitro* ont permis de caractériser les facteurs cellulaires nécessaires à l'activité des séquences IRES, et selon leur type les IRES recrutent des facteurs d'initiation différents (Pestova et Hellen, 2003). D'autres facteurs cellulaires (n'appartenant pas au groupe des facteurs d'initiation) sont également nécessaires à l'activité des séquences IRES. Il s'agit de facteurs ayant une activité *trans*-activatrice sur les IRES, regroupés sous le nom d'ITAF (IRES *Trans*-Activating Factor).

VII-2-B Les facteurs cellulaires interagissant avec les IRES viraux

Plusieurs équipes ont focalisé leurs recherches sur l'établissement d'une nomenclature simple et précise permettant de rassembler tous les IRES. Cependant, cet exercice a toujours été difficile à cause de la complexité du système. Bien que possédant la capacité d'initier la traduction, les IRES nécessitent un certain nombre de facteurs cellulaires. Deux types de facteurs sont nécessaires à l'initiation de la traduction, des facteurs canoniques (indispensables à l'initiation de la traduction Coiffe dépendante) et des facteurs ne faisant pas partie de cet ensemble et nommés non canoniques ou encore ITAF pour IRES *Trans* Activating Factors (Jackson et Kaminski 1995).

VII-2-B-1 Les facteurs non canoniques :

Il existe un grand nombre de facteurs rassemblés sous cette appellation ITAF, mais ils possèdent tous la même caractéristique : ils sont impliqués dans la régulation des facteurs canoniques et ne permettent pas de recruter directement les sous-unités ribosomiques.

PTB ou « Polypyrimidine-Tract Binding Protein » fait partie du complexe d'épissage de l'ARN dans la cellule non infectée. Cet ITAF stimule significativement les IRES de type I (Borman *et al.* 1993; Hellen et de Breyne 2007) et peut se lier passivement au type II (Kolupaeva *et al.* 1996) mais son importance fonctionnelle est encore controversée. Il fixe plusieurs sites dans les IRES du *Poliiovirus* (Hellen *et al.* 1994) par l'action des quatre motifs MRR (Perez *et al.* 1997), ce qui pourrait promouvoir ainsi leur bon repliement à la manière d'une chaperonne.

La protéine **La** pour « Lupus autoantigen » (Meerovitch *et al.* 1989) est, quant à elle, impliquée dans la traduction de l'ARNm, en se fixant à l'extrémité 3' de son IRES, permettant ainsi le choix correct du codon AUG initiateur (Svitkin *et al.* 1994; Ali et Siddiqui 1997).

hnRNP-L pour « heterogeneous nuclear RiboNucleoProtein L » est aussi capable d'interagir avec l'IRES du VHC (Hahm *et al.* 1998). Il se fixerait en 3' et augmenterait le taux de traduction des protéines virales.

ITAF₄₅ pour « IRES *Trans-acting* Factor » 45kDa est une protéine du cycle cellulaire exprimée exclusivement par des cellules en prolifération et stimule l'IRES du FMDV et coopère avec PTB (Pilipenko *et al.* 2000).

PCBP-2 pour « Poly(C) Binding Protein 2 », appartient à la famille des hnRNP. Ce facteur cytoplasmique a la capacité de se fixer à l'ARN et plus particulièrement à tous les IRES de la famille des *Picornavirus*. Il est important de noter que ce facteur est sans effet sur les IRES d'EMCV ou du FMDV (Graff *et al.* 1998). Son interaction avec le VHC a été établie mais requiert une étape d'activation de l'IRES, ce qui laisse supposer une reconnaissance tridimensionnelle plutôt qu'une séquence spécifique (Spangberg et Schwartz 1999).

Unr pour « Upstream of N-ras protein » stimulerait l'activité IRES (Hunt *et al.* 1999) en augmentant spécifiquement l'affinité de l'IRES du *Rhinovirus* pour PTB.

Il est étonnant de constater qu'aucun de ces facteurs n'a réellement un rôle fonctionnel concernant l'initiation de la traduction. Tous ces facteurs interagissent avec les différents IRES et ne sont pas responsables du recrutement direct de la sous-unité ribosomique 40S. Leurs actions seraient comparées aux chaperonnes en favorisant le bon repliement des structures présentes dans les IRES.

VII-2-B-2 IRES indépendantes des facteurs d'initiation canoniques :

D'autres IRES viraux sont au contraire, largement indépendants des facteurs d'initiation canoniques, et en particulier du «complexe eIF4F». Le groupe des Flaviviridae (VHC par exemple) possède des besoins en facteurs canoniques très différents de ceux des *Picornavirus*. L'IRES du VHC possède la capacité de recruter indépendamment la plupart des facteurs, la sous-unité ribosomique 40S et le complexe protéique eIF3. Le recrutement de la sous-unité 40S est effectué par l'interaction d'au moins deux protéines ribosomiques, S9 (Pestova *et al.* 1998) et S5 (Fukushi *et al.* 2001) avec différents domaines de l'IRES (Kolupaeva *et al.* 2000). L'IRES recrute via la partie apicale du domaine III le complexe eIF3 (Sizova *et al.* 1998) qui impose la stabilisation du complexe IRES/40S (Kieft *et al.* 2001; Kieft *et al.* 2008). Cette propriété de recruter la sous-unité 40S, indépendamment de la plupart des facteurs canoniques, est aussi retrouvée pour les IRES CrPV. Cette particularité confère à ces virus un avantage dans la rapidité de production de leurs protéines.

VII-3 Définition des types d'IRES

Comme il a été précisé plus haut, les IRES ont besoin de ces facteurs pour diriger activement l'initiation interne de la traduction. C'est dans cette optique que de nombreuses équipes se sont consacrées à l'étude du rôle des facteurs canoniques impliqués dans l'initiation de la traduction coiffe indépendante. Anciennement, la classification des groupes d'IRES était liée à la structure secondaire. La nomenclature actuelle proposée par Kieft est basée sur le recrutement des différents facteurs cellulaires. Suite à cette analyse, il est alors possible de les ordonner en fonction de la classe et du nombre de facteurs recrutés, pour une revue (Kieft *et al.* 2008). On peut en effet comparer l'IRES du PV qui nécessite un grand nombre de facteurs, à celui de l'IRES du virus CrPV qui ne nécessite aucun facteur pour recruter la sous-unité 40S. La figure ci-contre illustre le fonctionnement théorique des quatre groupes d'IRES (**FIG16**) (Kieft *et al.* 2008; Balvay *et al.* 2009).

Groupe I : C'est le système le plus simple jamais décrit. L'IRES permet de recruter directement la sous-unité ribosomique 40S sans la présence d'autres facteurs.

Groupe II : L'IRES permet de recruter les sous-unités 40S mais nécessite la présence des facteurs canoniques eIF tels que eIF3 et eIF2.

Groupe III : L'IRES nécessite la présence d'un grand nombre de facteurs canoniques ainsi que la présence de facteurs ITAF (IRES *Trans* Acting Factor) pour initier la traduction.

Groupe IV : Il est identique au groupe III, mais le codon initiateur AUG est situé plus en amont de l'IRES.

Afin de compléter la présentation de ces différents IRES, il est important de porter un regard précis sur les facteurs cellulaires déterminant l'initiation de la traduction de manière Coiffe indépendante.

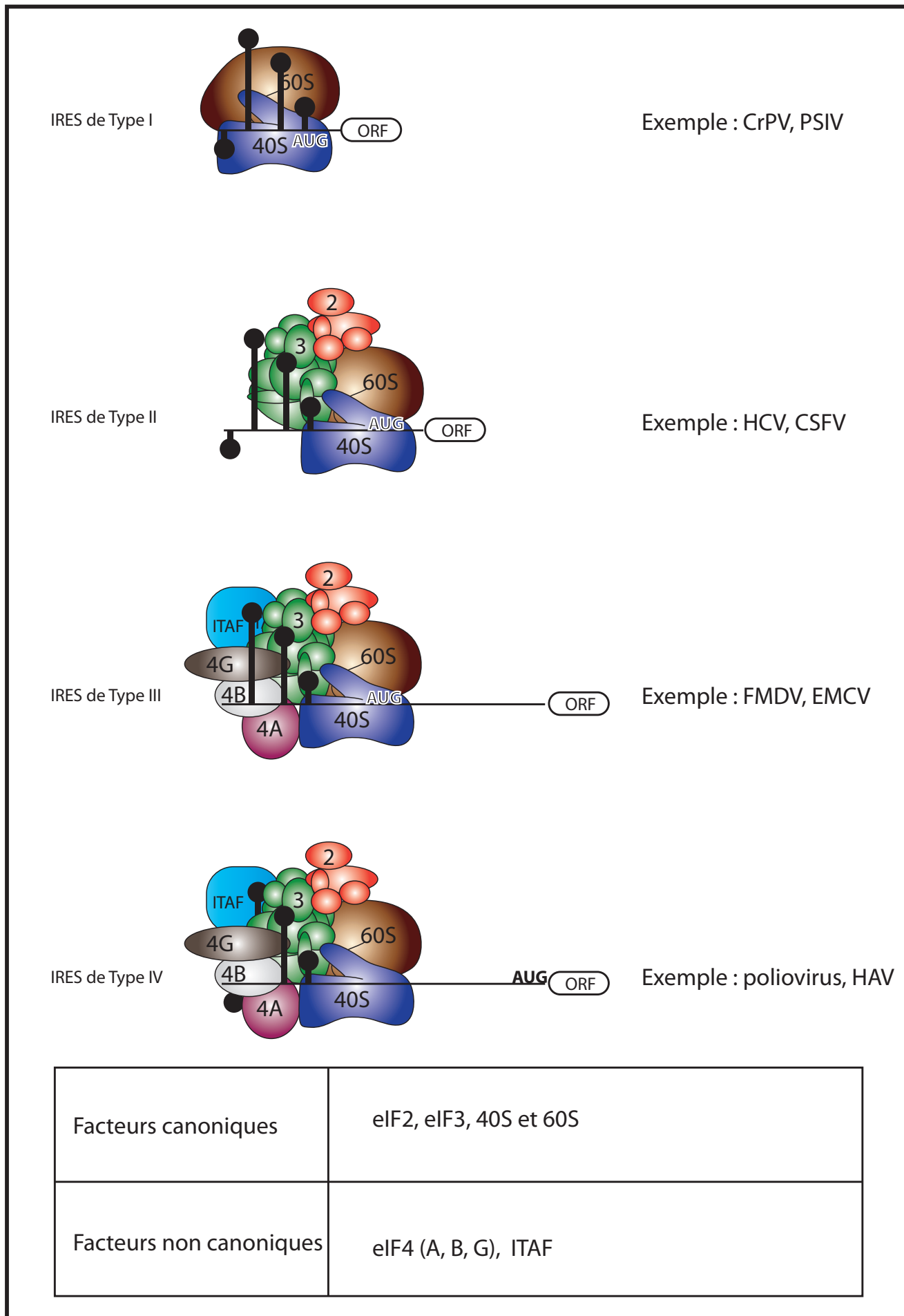


FIG16 : Diagramme simplifié représentant les 4 types d'IRES ainsi que leur besoin en facteurs (canoniques et non canoniques) dans la voie de l'initiation de la traduction CAP indépendante.

VIII L'IRES du VHC : étude structurale et fonctionnelle

VIII-1 Architecture de l'ARN viral

La région 5'UTR du génome de l'hépatite C (372nt) possède deux fonctions : une séquence IRES (39-372nt), impliquée dans l'initiation de la traduction, et un site de reconnaissance d'un microRNA hautement exprimé dans les cellules hépatiques : le miR-122 à la position 21 à 28nt de l'IRES (Jopling *et al.* 2005; Lancaster *et al.* 2006). De nombreuses études structurales et biochimiques ont été menées par plusieurs équipes en vue d'obtenir le plus de détails possible sur les mécanismes impliqués dans l'initiation de la traduction coiffe indépendante. C'est en 1992 que les équipes de Nomoto et Lloyd (Tsukiyama-Kohara *et al.* 1992; Byrd *et al.* 2005) ont montré que l'extrémité 5' UTR contenait une région très structurée et hautement conservée dans différents génotypes, pouvant correspondre à une séquence IRES. Plusieurs équipes ont confirmées par la suite la présence de cette structure, en se basant essentiellement sur les travaux effectués sur les IRES des *Picornavirus*. En 1999, l'équipe de Lemon a défini les bornes moléculaires minimales de cette séquence entre les nucléotides 40 et 341 de la région 5' de l'ARN (Honda *et al.* 1996a; Honda *et al.* 1996b; Honda *et al.* 1999; Rijnbrand *et al.* 2000). Différents travaux ont confirmés que cette structure était capable de se replier de façon autonome. Il adopterait alors une structure étendue, en présentant à la surface des sites de fixation à la machinerie de traduction (Kieft *et al.* 1999; Melcher *et al.* 2003). L'IRES du VHC est composé de quatre domaines (I à IV) très structurés dont trois sont essentiels pour l'activité IRES (II à IV). Des travaux de microscopie électronique ont permis de confirmer cette hypothèse : la structure tertiaire de l'IRES se composerait de domaines indépendants de longueur variables (Beales *et al.* 2001) (**FIG17**).

Le domaine I (noir) est situé à l'extrémité 5' et forme une épingle à cheveux de 27 nt. Celui-ci pourrait jouer un rôle en régulant la traduction des protéines virales de manière négative (Han, J.H *et al.* 1991). Des structures de ce type sont présentes dans certains ARNm et pourraient favoriser l'étape de scanning du ribosome.

Le domaine II (rouge) est le domaine le plus étudié de l'IRES du VHC et de nombreuses structure tridimensionnelles sont disponibles. Il est composé d'une boucle dans la partie basale (IIa) suivit d'un motif E-loop formant une courbure qui se prolonge jusqu'à une tige/boucle apicale (IIb). La structure de ce domaine a été résolue par la technique de RMN qui indique que ce domaine est peu flexible et composé de deux domaine IIa et IIb (Brown *et al.* 1992; Hellen *et de* Breyne 2007) (**FIG17**).

Le domaine III (vert) est la plus grande région formant l'IRES du VHC. Il comporte une succession de tiges/boucles en épingle à cheveux organisée en jonction avec 3 ou 4 brins d'ARN (IIIabcdef) (Brown *et al.* 1992). On peut dissocier le domaine III en trois parties. La première partie rassemble la région apicale IIIabc qui se compose d'une jonction 3 brins (IIIabcJunc),

dont la structure a été résolue par cristallographie (Kieft *et al.* 2002). Cette structure originale présente une jonction en «X» qui fait intervenir deux hélices en parallèle liées par un simple croisement sous l'influence d'une liaison phosphate entre IIIa et IIIb. Cette jonction est impliquée dans la fixation du facteur eIF3. La partie centrale du domaine III est représentée par le sous-domaine IIIId qui est le plus conservé des IRES de type « HCV » (Brown *et al.* 1992). Cette tige boucle présente une structure en E-loop asymétrique, pouvant être assimilée à un ARN ribosomal. Ce domaine expose une tétraboucle UGGG impliquée dans la liaison avec la sous-unité 40S. La dernière partie du domaine III est composée des sous-domaines IIIe et IIIf en formant une jonction 4 brins dont la structure n'a pas encore été résolue. Ce dernier motif a été retrouvé dans un grand nombre d'IRES et semble présenter des motifs clefs pour l'initiation (Hellen et de Breyne 2007). La boucle IIIe expose quant à elle des nucléotides en formant une boucle rigidifiant la structure du domaine III (Rijnbret et Lemon 2000). Le domaine IIIf forme lui un pseudo noeud dont la structure n'est pas connue (**FIG17**).

Le domaine IV (en bleu) est la partie basale de l'IRES. Ce domaine n'est pas conservé parmi les IRES et nous ne disposons pas de données structurales le concernant. Son architecture présente une tige/boucle (Brown *et al.* 1992; Honda, Brown *et al.* 1996a; Hellen et de Breyne 2007) (**FIG17**).

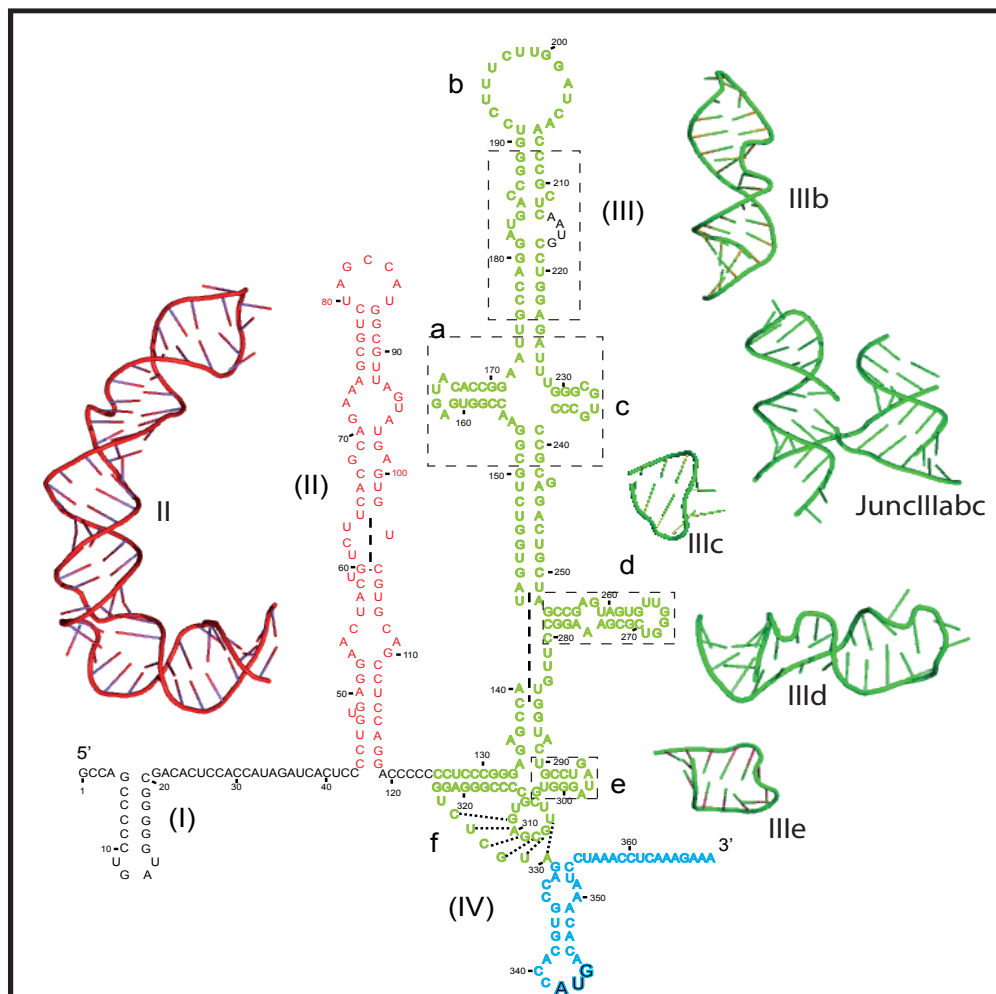


FIG17 : Information structurale de l'IRES et présentation des zones d'interactions. Source : Lukavsky et al. 2009

Comme nous pouvons le constater, l'IRES du VHC fait l'objet de plusieurs études structurales. L'étude de ces différents fragments a permis d'appréhender l'architecture de ces motifs et de comprendre leurs arrangements moléculaires. Cependant, l'étude de ces domaines individualisés ne permet pas de reconstituer la structure de l'IRES complet (seul 50% de la structure est connue). Les techniques structurales actuelles ne sont pas encore adaptées à l'étude de molécule d'ARN d'une telle longueur. C'est dans ce but que plusieurs équipes ont utilisé la microscopie électronique afin d'étudier l'IRES en présence de différents facteurs protéiques comme par exemple la sous-unité 40S. Ces études ont permis de mieux comprendre l'arrangement des domaines de l'IRES en présence de protéines. Néanmoins, les informations concernant l'IRES VHC - seul en solution - ne sont pas disponibles actuellement. Ces données permettraient de faire la lumière sur les mécanismes conduisant à la formation des structures nécessaires au recrutement des différents facteurs permettant l'initiation de la traduction de manière indépendante de la coiffe.

VIII-2 L'IRES du VHC est les facteurs de l'initiation coiffe indépendante

L'IRES du VHC est impliqué dans la capture d'une partie de la machinerie de traduction eucaryote (eIF3, eIF2, 40S et 60S) et peut être comparé de ce fait à un mécanisme procaryote en remplaçant la séquence Shine-Dalgarno qui est capable de recruter la petite sous unité ribosomique 30S (Shine et Dalgarno 1975). L'IRES du VHC peut recruter directement la sous-unité 40S par l'intermédiaire du domaine III (Honda *et al.* 1999; Kieft *et al.* 2001; Otto *et al.* 2004) (**FIG18**). Concernant le domaine IV, celui-ci semble posséder un rôle majeur dans l'activité IRES car la plupart des mutations touchant le domaine IV affectent radicalement cette activité, mais ne semble pas diminuer le recrutement de 40S (Kolupaeva, Pestova *et al.* 2000). Le domaine II quant à lui ne semble pas directement impliqué dans le recrutement de la sous-unité 40S (Kieft *et al.* 2001; Otto *et al.* 2004; Lukavsky 2008). La formation de ce pré-complexe est fortement stabilisée par le recrutement du facteur eIF3 et du complexe eIF2-ARNt-Met (Fraser *et al.* 2004). La formation du complexe 80S nécessite la présence du facteur 60S et de l'hydrolyse du GTP et requiert les domaines II et III (Fraser *et al.* 2004; Otto *et al.* 2004; Boehringer *et al.* 2005). Comme évoqué plus haut, la plupart des mutations visant le domaine III influe négativement sur la formation des pré-complexes, inhibant ainsi la formation du complexe 80S (Ji *et al.* 2004). Concernant le domaine II, celui-ci a un rôle majeur dans le positionnement du complexe 80S, et son absence conduit à une diminution importante de l'activité translationnelle (Ji, Fraser *et al.* 2004; Otto *et al.* 2004; Locker *et al.* 2007). D'après ces informations on peut constater que l'IRES est agencé en 3 domaines distincts possédant des fonctions complémentaires dans l'initiation de la traduction. Certains domaines vont recruter des facteurs, d'autres vont stabiliser la structure. L'intégralité de la séquence IRES ayant été défini comme nécessaire à l'activité IRES on comprend mieux que chaque domaine qui possède une fonction

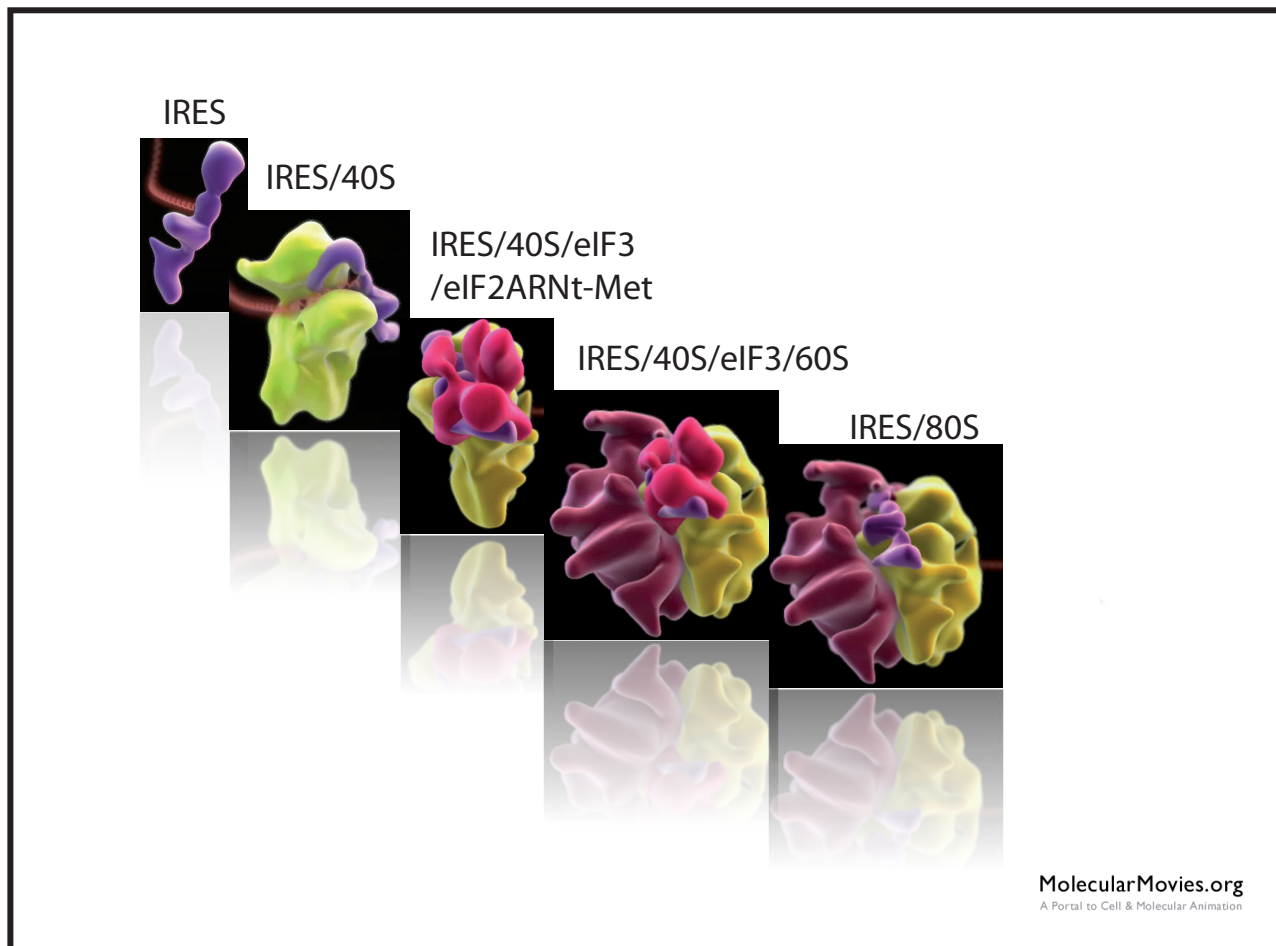


FIG18 : Modèle en 3D de l'initiation de la traduction dans le cas du VHC par l'IRES conduisant à la formation du complexe 80S.

Source : Molecular movies - <http://www.molecularmovies.com>

propre doit se retrouver dans une conformation particulière pour permettre l'initiation de la traduction. Plusieurs résultats présents dans la littérature convergent vers la même hypothèse qui indique que la séquence IRES nécessite une étape d'activation avant d'initier réellement la traduction coiffe indépendante.

IX Le complexe eIF3/IRES

Plusieurs études ont démontré que l'IRES du VHC interagit fortement avec le facteur eIF3 (Sizova, Kolupaeva *et al.* 1998; Kieft *et al.* 1999; Siridechadilok *et al.* 2005). Plusieurs travaux présentent différentes sous-unités du complexe eIF3 comme impliquées dans l'interaction avec l'IRES. En effet les sous-unités eIF3a, b, c, d et f sont retrouvées en présence de l'IRES (Zhou *et al.* 2008). En utilisant la spectrométrie de masse, les auteurs ont pu identifier d'une part l'arrangement des différentes sous-unités du complexe eIF3 et d'autre part d'identifier 5 sous-unités permettant le recrutement de l'IRES du VHC. Ces travaux confirment les précédents résultats obtenus par étude *in vitro* de «pontage» chimique ou aux UV indiquant que ces cinq sous-unités présentaient des activités de liaison à l'ARN (Sizova *et al.* 1998). La présence de motif MRR dans deux sous-unités, eIF3b et eIF3g nous indique le potentiel rôle de ces pro-

téines dans la formation du complexe d'initiation. Cependant, seul le rôle dans l'interaction avec 40S du domaine MRR de la protéine eIF3g a été caractérisé. Les informations concernant le complexe eIF3/IRES sont assez fragmentaires. Nous disposons cependant d'une reconstitution en 3D du complexe eIF3/IRES générée par microscopie électronique (Siridechadilok *et al.* 2005). Les résultats indiquent que l'IRES posséderait une forme allongée (250 à 300Å) comme cela est retrouvé en présence de 40S. Cependant, les auteurs n'ont pas pu identifier clairement les domaines de l'IRES impliqués dans l'interaction. Les données de la littérature nous informent que le domaine apical de l'IRES (IIIabc) est protégé des RNases en présence du complexe eIF3. D'après ces résultats, nous pouvons conclure que ce domaine est impliqué dans la formation du complexe. Cependant, il n'est pas possible de déterminer avec exactitude quels sont les sous-unités du complexe eIF3 responsables de cette interaction. De ce fait, plusieurs questions se posent : La protéine eIF3b est-elle impliquée ? son domaine MRR est-il capable de lier l'ARN viral ? Le domaine IIIabc est-il le seul domaine capable d'interagir avec le complexe eIF3 ? L'IRES est-il capable de se lier directement au complexe eIF3 ou doit-il changer de conformation pour se lier efficacement ?

X Objectifs de thèse

L'objectif de mon travail de thèse consistait à élucider, par une approche structurale et biochimique, les mécanismes impliqués dans la formation du complexe eIF3/IRES dans le cadre de l'initiation de la traduction coiffe indépendante du VHC. Il a été démontré par d'anciennes techniques que cinq sous-unités du complexe eIF3 interagissaient avec l'IRES du VHC. Cependant, les informations concernant les sous-unités impliquées ne sont que fragmentaires et ne permettent pas une étude fine du complexe. Notamment, le rôle de la protéine eIF3b et de son motif MRR n'a pas été testé. De plus, la protéine eIF3b n'a pas encore fait l'objet d'une quelconque étude biophysique et structurale. Cette protéine joue pourtant un rôle clef dans la structuration du complexe eIF3 et semble interagir avec l'IRES du VHC. Les informations structurales concernant la séquence entière de l'IRES en solution sont également absentes de la littérature. Afin d'étudier cet ARN, j'ai conduit plusieurs études biophysiques et structurales pour étudier ce fragment important du génome viral.

Le premier chapitre concerne l'étude structurale de l'ARN viral en solution qui est étudié par la technique de SAXS. Grâce à cette technique, j'ai été en mesure de modéliser la forme des différents domaines de l'IRES et de reconstruire l'enveloppe de l'ARN viral. La technique de SAXS étant essentiellement dédiée à l'étude de protéine, il a été nécessaire d'adapter les techniques de modélisations afin d'étudier l'enveloppe de cet ARN en solution.

Le deuxième chapitre est dédié à l'étude de la protéine eIF3b et de son motif MRR. Les données biochimiques et structurales étant peu présentes dans la littérature j'ai initié diffé-

rentes approches qui m'ont conduit également à utiliser la technique de SAXS.

Enfin, le troisième chapitre aborde l'étude de l'interaction de l'IRES du VHC avec la protéine eIF3b. Plus précisément, j'ai étudié, par différentes techniques biochimiques, l'interaction entre l'IRES du VHC et le domaine MRR de la sous-unité b du complexe eIF3.

L'ensemble des résultats présentés ici sont le fruit de trois années de recherche menée à l'UVHCI sous la direction du Pr. Emmanuel Drouet et du Dr. Florence Baudin. Au cours de mes recherches, j'ai été amené à collaborer avec le Pr. Marc Jamin (UVHCI) pour le traitement des données de SAXS, avec le Dr. Jan Medenbach (EMBL Heidelberg) pour les échantillons de complexe eIF3 ainsi qu'avec le Dr. Jérôme Boisbouvier (IBS) pour l'étude RMN du complexe MRR-eIF3b/IRES.

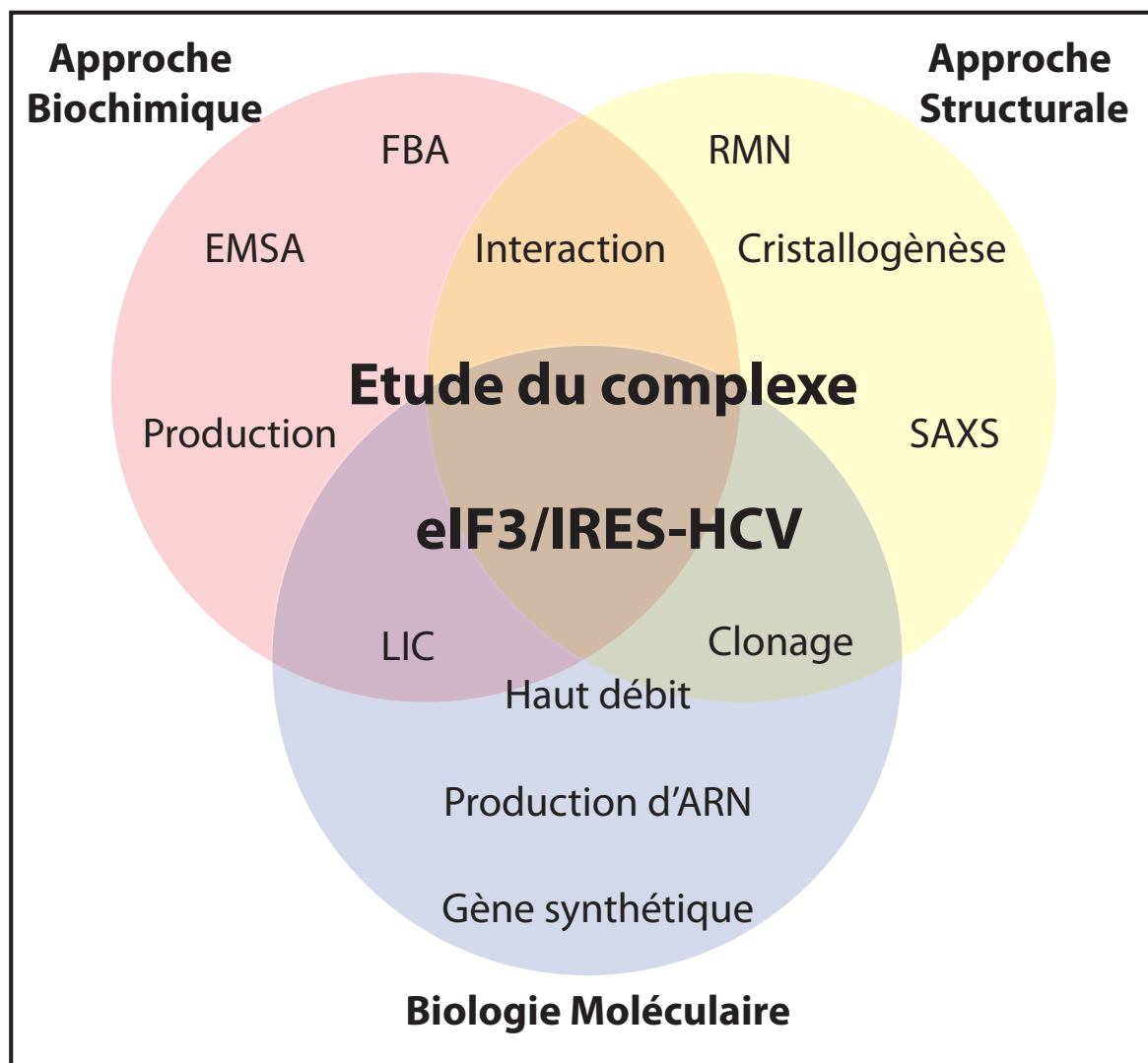


FIG19 : Diagramme des différentes approches techniques utilisées pour l'étude structurale et fonctionnelle du complexe eIF3/IRES-HCV

MATERIEL & METHODES

I Etude des protéines du complexe eIF3 humain

Nous avons utilisé deux systèmes de clonage pour étudier et exprimer ces protéines chez la bactérie : une approche classique (eIF3b, les MRR et eIF3j) et une approche haut débit par utilisation de la technique de LIC ou ligation Independent Cloning (les sous-unités du complexe eIF3). Les gènes codant pour les protéines d'intérêt proviennent de la banque d'ADNc de la société ImaGene, du laboratoire du Dr Kazusa ainsi que du laboratoire du Dr J.Hershey. La liste des protéines étudiées est présentée dans le **Tableau VII**.

Tableau VII : Liste des constructions réalisées pour l'étude

Nom	Position	MW kDa
eIF3b	814 (MQDA-GNQE)	93
	682 (GRAA-GNQE)	80
	539 (SDEW-GNQE)	63
	113 (LLWR-GNQE)	14
	426 (SDEW-LLWR)	50
	133 (MQDA-GRAA)	13
	569 (GRAA-LLWR)	65
	144 (GRAA-SDEW)	16
	630 (SVIV-GNQE)	74
	702 (MQDA-LLWR)	80
	276 (MDQA-SDEW)	30
	96 (SVIV-SDEW)	11
	185 (MQDA-SVIV)	18
	83 (SVIV-LFTD)	10
MRR-eIF3g	79 (ATIR-WAKP)	9
eIF3j	251 (AGDS-EDFM)	28
eIF3a	1382 (MPAYF-TTVRR)	166
	880 (MPAYF-RRLGD)	105
	180 (LEEL-REEE)	23
	522 (DSSL-TTVRR)	63

Nom	Position	MW kDa
eIF3c	913 (MPRF-QTAY)	105
	371 (PKRV-FANP)	42
	436 (RCIL-DHKQ)	51
	558 (RGTEI-DHKQ)	70
eIF3d	532 (LAKF-SDED)	62
	164 (LAKF-VEVR)	20
	360 (RSDW-LPDG)	41
eIF3e	445 (LAEY-SGFY)	52
	380 (LAEY-NARL)	45
eIF3f	357 (MATP-LVNL)	38
	268 (RVVR-LVNL)	30
eIF3h	352 (MASR-EYNN)	40
	314 (QVQI-EYNN)	36
eIF3i	320 (MKPI-EFEA)	36
eIF3E1	325 (MSYP-GQRP)	37
	295 (PKQD-GQRP)	32
eIF3M	531 (MSVP-LSDT)	63
	290 (DLAQ-QTQR)	33

Chaque cible a été clonée dans trois vecteurs d'expression (**Tableau VIII**). Ces trois vecteurs nous permettent de tester l'effet de trois étiquettes (6xHistidine, GST ou MBP) sur la solubilité des protéines.

Tableau VIII : Liste des vecteurs utilisés pour le clonage et l'expression des gènes

Nom du vecteur	Cassette de résistance aux antibiotiques	Etiquette	Site de restriction enzymatique utilisé pour le clonage classique
pET-TEV	Kanamycine	poly-his clivable TEV	NdeI/XhoI ou NcoI/XhoI
pGEX-TEV	Ampicilline	GST clivable TEV	EcoRI/BamHI
pETM40	Kanamycine	MPB clivable TEV	NdeI/XhoI

Les cartes des vecteurs ainsi que les oligonucléotides générés pour ces clonages sont présentés dans l'**Annexe II** (page 184).

I-1 Biologie moléculaire

I-1-A Clonage des protéines eIF3b, eIF3b, eIF3J, MRR-eIF3b et MRR-eIF3g

Les gènes codant pour chaque protéine ont été amplifiés par PCR. La bande d'ADN d'intérêt issu du produit de PCR est purifiée par extraction sur gel à la suite d'une électrophorèse (kit de purification NucleoSpin Extract II, société Macherey Nalgen). L'insert ainsi purifié est ensuite coupé en parallèle avec les vecteurs cibles par les couples d'enzymes de restriction appropriées (**Tableau VIII**). Après cette étape, les produits sont de nouveau purifiés et concentrés par évaporation avant d'être utilisés pour la ligation. La quantité d'ADN est déterminée par spectroscopie d'absorbance à 260nm. La ligation est réalisée à l'aide de la T4 DNA ligase en présence d'ATP avec une concentration 3 fois plus élevée en insert qu'en vecteur. Après une incubation à 37°C pendant 30 minutes, le produit de ligation est ensuite introduit par transformation dans des bactéries *E.coli* K12 TOP10 (Stratagen) avant d'être sélectionnées sur boîte de pétri. Plusieurs clones de chaque construction, font l'objet d'une culture liquide en petit volume afin de produire de petites quantités d'ADN (kit MINIPREP de la société QIAGEN). Chaque construction est ensuite séquencée par la société MWG.

I-1-B Clonage par la technique LIC (Ligation Independant Cloning) des protéines du complexe eIF3.

La technique LIC a été utilisée pour obtenir des sous-domaines de la protéine eIF3b (12 constructions) ainsi que pour les différentes sous-unités du complexe eIF3 dans les 3 vecteurs présentés dans le **Tableau VIII**. La technique LIC consiste à générer des fragments simple brin d'une longueur de 12 à 15 nt de part et d'autre de l'insert et du vecteur soit une séquence LIC1 en 3' et une séquence LIC2 en 5'. La présence de ces séquences simple brin est suffisante pour favoriser la fusion de l'insert dans le vecteur comportant les séquences LIC1 et 2 complémentaire. Afin de créer les séquences LIC simple brin de part et d'autre de l'insert et aux extrémités du vecteur, il est nécessaire d'utiliser les propriétés de la T4 DNA polymérase en présence de nucléotide. Cette enzyme possède une activité exo nucléase 3', qui va éliminer les nucléotides et généré une séquence simple brin. La séquence LIC1 est constitué des nucléotides A, G, T sur 15 nucléotides et comporte un nucléotides C à la position 16. L'enzyme va donc éliminer les nucléotides A, G et T et s'arrêtera à la position 16 du fait de la présence de dCTP dans le milieu réactionnelle. La réaction finale de ligation est réalisée par la bactérie qui jouera le rôle de la T4DNA ligase. Le vecteur étant linéaire et comportant deux séquences différentes ne peut se re-circulariser. La présence de clones sur boîte de pétri après la réaction est synonyme de la présence de la construction à 99%.Cependant pour plus de sureté chaque construction fera l'objet d'un séquençage.

Préparation des vecteurs et des inserts (création de la séquence LIC)

Une PCR est réalisée sur le vecteur et sur le gène d'intérêt avec les oligonucléotiques comportant les séquences LIC1 et LIC2 (**Annexe II** page 184).

Protocole pour la réalisation d'une PCR

Oligo LIC1	1 µl à 10pmol/µl
Oligo LIC2	1 µl à 10pmol/µl
H ₂ O qsp 20µl	5,4 µl
ADN	2 µl à 50ng/µl
DMSO 100%	0,6 µl soit 3% finale
Master mix 2x Phusion GC rich (Finnzyme)	10 µl

Programme utilisé pour la réalisation de l'étape de PCR

Cycle 1 (x3)		Cycle 2 (x30)		Cycle 3 (x1)	
96°C	2 min	96°C	2 min	72°C	10 min
72°C	1min	63-68°C	1min + 1s par cycle	15°C	sur la nuit

Après la réaction, le produit de PCR est purifié par électrophorèse sur gel d'agarose (Kit NucleoSpin extract II, Macherey Nalgen). Une fois le vecteur purifié, celui-ci subit une étape de coupure à l'endonucléase DpnI (NEB) afin d'éliminer le brin d'ADN matrice (éventuellement présent) qui est encore méthylé. Cette étape est importante et permet de garantir la présence du vecteur amplifié comportant les séquences LIC1 et LIC2. Après cette étape, l'enzyme est inactivée à la chaleur (80°C pendant 20 min). Les ADN purifiés sont additionnés de 2 unités de T4 DNA polymérase (NEB) dans le tampon T4LIC : 33 mM Tris-acetate pH 8,0; 66 mM d'AcK; 10 mM d'AcMg; 0,5 mM DTT et 100 µg/ml de BSA en présence de 0,5mM de dCTP (Vecteur), dGTP (Insert). Le mélange réactionnel est incubé pendant 30 minutes à 37°C avant d'être inactivé 10 minutes à 70°C. Le vecteur et l'insert sont ensuite une dernière fois purifiés par électrophorèse d'agarose avant d'être concentrés à 100ng/µl par évaporation. Le vecteur dispose donc d'une séquence LIC1 simple brin en 3' composée des nucléotides A, G et T qui est complémentaire à la séquence LIC1 simple brin présent en 3' de l'insert composé de A, C et T. Il est en de même pour l'extrémité 5' qui comporte la séquence LIC2. Après une estimation de la quantité d'ADN par spectroscopie d'absorbance à 260nm, 75ng d'insert sont mélangés à 25ng de vecteur. L'étape de fusion est réalisée dans un tampon LIC 25mM tris-HCl pH 7,8 ; 10mM MgCl₂ ; 4mM βME ; 0,4mM ATP. Après une incubation de 30 minutes à température ambiante, le mélange réactionnel est introduit par transformation dans des bactéries Top10 (Stratagene).

I-2 Biochimie des protéines

I-2-A Expression et purification : *eIF3b-MRR*, *eIF3g-MRR*, *eIF3j*

Sur-expression des protéines :

La sur-expression des protéines est réalisée dans la bactérie BL21 (DE3). Les cultures se font sous agitation à 200 rpm à 37°C dans 1 litre de LB (Luria Broth) contenant l'antibiotique souhaité (Carbenicilline 100µg/ml ou Kanamycine 60µg/ml). L'induction de l'expression de la protéine se fait à une DO_{600nm} de 0,6 avec de l'IPTG 0,5 mM final pendant 4-5 heures à plusieurs températures (15, 25 ou 37°C) afin de tester la solubilité des protéines. Les bactéries sont ensuite centrifugées à 10000 g pendant 10 minutes à 4°C et les culots sont congelés à -80°C.

Purification:

Le culot de bactérie est décongelé rapidement à 37°C avant d'être repris par un tampon de lyse (20ml de tampon A pour 4g de pâte cellulaire soit environ 1L de culture). Tampon A : 20mM Tris HCl pH 7,5; 150 mM NaCl; 10mM imidazole fraîchement préparé (SIGMA), 1 tablette d'anti protéase Complete EDTA free (Roche). La lyse est réalisée par sonication 3x5 minutes à 80% d'intensité avec une pulsation de 0,5s en présence de lysozyme à 1µg/ml à 4°C. Le lysat est ensuite centrifugé à 45000 g en JA 25.50 à 4°C pendant 45 minutes. Le surnageant est ensuite déposé sur une résine d'affinité (Nickel : His Select Cobalt (SIGMA)) préalablement équilibrée en tampon A. La résine est ensuite lavée par 2 volumes de colonne en tampon A avant d'être lavée par 2 volumes de tampon B = tampon A + 1M NaCl. Ce lavage supplémentaire a pour but d'éliminer les acides nucléiques résiduels. La colonne est de nouveau lavée en tampon A avant l'étape d'élution. L'élution est réalisée en palier d'imidazole (10, 100, 200, 300, 400 et 500mM) : Tampon C = tampon A + 10 à 500mM imidazole. Les fractions sont ensuite analysées sur gel SDS PAGE 15%. Les fractions d'intérêts sont rassemblées et concentrées sur membrane Millipore ULTRA 10kDa jusqu'à obtention d'un volume final de 5ml.

Une étape supplémentaire peut être réalisée afin d'éliminer les étiquettes encore présentes grâce à l'action de la protéase TEV. Tout d'abord une étape de dialyse est réalisée afin d'éliminer l'imidazole avant d'effectuer la digestion (1µg de TEV pour 100µg de protéine) à température ambiante toute la nuit. Le produit de la digestion est ensuite déposé sur une résine Ni 2+ équilibrée dans le tampon A puis la fraction non retenue contenant la protéine clivée est conservée. Afin d'augmenter le degré de pureté des protéines, une étape de chromatographie d'exclusion de taille est réalisée à 4°C. Les échantillons sont tout d'abord centrifugés 10 minutes à 25000g avant d'être injectés sur la colonne Superdex 75 HiLoad 16/60 (GE Healthcare) préalablement équilibrée en tampon A. Toutes les fractions sont collectées et sont analysées sur gel SDS PAGE 15%. Les fractions d'intérêts sont rassemblées et conservées à 4°C avant d'être concentrées avec une unité de concentration Millipore ULTRA de 5kDa. Les protéines sont ensuite con-

gelées en azote liquide en présence de 5% de glycérol final et conservées à -80°C.

I-2-B Expression et purification du domaine eIF3b-MRR pour l'étude RMN

Ce protocole a été développé et adapté par Isabel AYALA (laboratoire de marquage LRMN de l'IBS). L'obtention d'une protéine marquée 15N ou 15N/13C nécessite une gestion du temps différente de la culture classique.

La transformation des bactéries BL21(DE3) avec le plasmide d'intérêt est effectuée le jour 1. Le deuxième jour, la préculture en milieu LB est réalisée le matin, soit 1 clone dans 5ml de milieu additionné des antibiotiques, la culture est réalisée à 37°C à 200rpm. Après 10h de culture, 1ml de cette première préculture est utilisé pour inoculer 100ml d'une seconde préculture en milieu M9 comportant l'isotope souhaité, la culture est réalisée sur la nuit à 37°C à 200rpm.

Le troisième jour, la préculture de la nuit est utilisée pour inoculer 1L de milieu M9. La préculture est tout d'abord centrifugée pendant 20 minutes à 4000g avant d'être reprise par le milieu de culture pour obtenir une DO_{600nm} avoisinant 0,2. La culture est incubée à 37°C à 200rpm jusqu'à obtention d'une DO_{600nm} de 0.6. La culture est ensuite additionnée de 1mM final d'IPTG et incubée toute la nuit à 18°C à 200rpm. Les cellules sont ensuite récoltées et traitées de la même manière qu'une culture en milieu normal et le culot bactérien est congelé à -80°C. Le protocole utilisé pour purifier la protéine est identique à celui utilisé pour la purification de la protéine non marquée.

I-2-C Expression et purification de la protéine eIF3b

Sur-expression de la protéine :

La sur-expression de eIF3b est réalisée dans la bactérie BL21RIL (DE3) (Stratagene) comportant un plasmide RIL produisant des ARNt supplémentaires pour les acides aminés R, I et L afin «d'humaniser» la bactérie. L'utilisation de plasmide RIL permet de favoriser la production de protéine comportant des codons. On retrouve généralement ces codons dans les protéines humaines. Les cultures se font sous agitation (200 rpm) à 37°C dans du milieu auto-induit AI (F. William Studier, Brookhaven National Laboratory, personal communication) contenant de la Kanamycine (60 µg/ml) final pendant 24 h.

Composition du milieu de culture AI (pour 1 litre) :

- 925ml de ZYP (5g d'extrait de levure et 10g de tryptone) autoclavé
- 1ml de $MgSO_4$ 1M autoclavé
- 20ml de solution 5052 (20x)

Glucose	25g/L
Glycérole	250g/L
Alpha-lactose	100g/L

-50ml de solution NPS (20x)

$(\text{NH}_4)_2\text{SO}_4$	0,5M
KH_2PO_4	1M
Na_2HPO_4	1M

Après la phase de culture, les bactéries sont ensuite centrifugées à 10000 g pendant 10 min à 4°C et les culots sont congelés à - 80°C.

Purification:

Le culot de bactérie est décongelé rapidement à 37°C avant d'être repris par un tampon de lyse (20ml de tampon A pour 4g de pâte cellulaire soit environ 1L de culture). Tampon D : 20mM HEPES pH 7,5; 150 mM NaCl; 10mM Imidazole fraîchement préparé (SIGMA), 1 tablette d'anti protéase Complete EDTA free (Roche). La lyse est réalisée par sonication 3x5 minutes à 80% d'intensité avec une pulsation de 0,5s en présence de lysozyme à 1µg/ml à 4°C. Le lysat est ensuite centrifugé à 45000 g en JA 25.50 à 4°C pendant 45 minutes. Le surnageant est ensuite déposé sur une résine d'affinité (Nickel : His Select Cobalt (SIGMA)). Chaque résine est préalablement équilibrée en tampon de lyse et la fraction non retenue est conservée à 4°C. La résine est ensuite lavée par 2 volumes de colonne en tampon D avant d'être lavée par 2 volumes de tampon E = tampon D + 1M NaCl. La colonne est de nouveau lavée en tampon D avant l'étape d'élution. L'élution est réalisée en palier d'imidazole (10, 100, 200, 300, 400 et 500mM) : Tampon F = tampon D + 10 à 500mM imidazole. Les fractions sont ensuite analysées sur gel SDS PAGE 10%. Les fractions d'intérêts sont rassemblées et concentrées sur membrane Millipore ULTRA 50kDa jusqu'à obtention d'un volume final de 5ml.

Toutes les étapes suivantes se déroulent à 4°C. Centrifugation de la protéine 10 minutes à 16000 g avant injection sur la résine Superdex 200 HiLoad 16/60 (GE Healthcare) préalablement équilibrée en tampon D. Après analyse sur gel SDS PAGE 10%, les fractions d'intérêt sont rassemblées et conservées à 4°C avant d'être concentrées jusqu'à 15 mg/ml (unité de concentration Millipore ULTRA 50kDa) . La protéine est ensuite congelée en azote liquide en présence de 5% de glycérol finale et conservée à -80°C.

I-2-D Purification du complexe eIF3

Le complexe eIF3 est purifié à partir de lysat de réticulocyte de lapin grâce à une collaboration avec l'équipe de Matthias Hentze de l'EMBL Heidelberg. Le protocole de purification du complexe eIF3 est en **Annexe III** (Page 188).

Elimination de la sous unité j du complexe eIF3 : eIF3Δj

Le facteur eIF3 purifié est déposé (100µl à 0,1µg/µl) sur un gradient de sucrose (10-30%) de 1 ml (tube de 1,2 ml) en tampon G : 400mM KCl; 20mM Tris-HCl pH 7,5; 2mM EDTA,

5mM DTT; 10mM benzamidine + 1 tablette d'anti protéase Complete EDTA free (Roche). Le complexe est centrifugé 22h à 45000 g, 4°C en SW55 + adaptateurs (BECKMAN). Après centrifugation, la sous-unité eIF3j est retrouvée en haut du gradient alors que le complexe eIF3Δj est situé au fond du gradient. Le complexe ainsi purifié est dialysé toute la nuit contre le tampon H 100mM KCl; 20mM Tris-HCl pH 7,5; 2mM EDTA; 5mM DTT; 10mM benzamidine + 1 tablette d'anti protéase Complete EDTA free (Roche). Le complexe est ensuite chargé sur une résine d'échange d'ion MonoQ (2 ml) préalablement équilibrée en tampon H monté sur un appareil de purification Biorad DuoFlow. Le complexe est purifié par un gradient linéaire de KCl (100 à 600mM) dans le tampon H. Les fractions sont rassemblées avant d'être chargées sur une résine de filtration sur gel superose 6 équilibrée en tampon H, montée sur le même appareil de purification. Les fractions sont concentrées et analysées sur gel SDS PAGE gradient 4-20% en acrylamide, avant d'être congelées en présence de 10% de glycérol final à -80°C.

Formation du complexe eIF3Δj/eIF3j recombinante :

Le complexe eIF3Δj/eIF3jrec est formé par mise en contact d'un excès de eIF3j recombinante et purifiée, en présence du complexe eIF3Δj. Le mélange est ensuite déposé sur un gradient de sucrose 10-30% en tampon H (tube de 1,2 ml). Le complexe est centrifugé 22h à 45000 g à 4°C en SW55. Après centrifugation, la sous unité eIF3jrec en excès est retrouvée en haut du gradient alors que le complexe eIF3Δj/eIF3jrec est retrouvé au fond du gradient. Le complexe est récupéré pour être chargé sur une résine d'exclusion de taille Superose 6 (GE Healthcare) préalablement équilibré en tampon H. Les fractions d'intérêts sont concentrées et analysées sur gel SDS PAGE en gradient 4-20% d'acrylamide avant d'être congelées en présence de 10% de glycérol final.

Expression et purification des protéines du complexe eIF3 :

La sur-expression des protéines a été réalisée dans la bactérie BL21 (DE3) RIL (Stratagene). Les cultures se font sous agitation (200 rpm) à 37°C dans du milieu AI contenant de la Kanamycine (60 µg/ml) final pendant 24 h en plaque 96 puits. Les cultures sont ensuite centrifugées avant d'être lysées chimiquement par l'utilisation de 100 µl de tampon BugBuster additionné de Lysozyme et de Benzonase (Novagen). Après l'étape de lyse, les échantillons sont centrifugés afin de séparer la phase soluble de la phase non soluble. Tous les échantillons sont analysés par immunodétection à l'aide de la technique dot blot. Après sélection des clones exprimant la construction souhaitée, la phase soluble est déposée sur un gel SDS PAGE. Les clones exprimant les protéines sous forme soluble et en grandes quantités sont alors sélectionnés et conservés en azote liquide en présence de 20% de glycérol final. Les échantillons solubles sont chargés sur les résines adéquates et traitées selon le protocole du fabricant. Chaque fraction est ensuite déposée sur une membrane de nitrocellulose afin de réaliser un nouveau dot blot permettant d'apprécier la présence de la protéine cible avant d'effectuer une

électrophorèse SDS PAGE. Les clones sélectionnés (bonne expression, soluble à plus de 50%) seront utilisés pour réaliser la production à plus grande échelle de chaque protéine.

Immunodétection

Préparation des anticorps pour utilisation de la technique de western blot ou dot blot.

Tableau IX : liste des anticorps utilisés pour l'immunodétection

Nom	Dilution	Détection
Anticorps anti Histidine (SIGMA)	1/5000	Couplé à HRP
Anticorps anti GST (SIGMA)	1/5000	Couplé à HRP
Anticorps anti MBP (SIGMA)	1/5000	Couplé à HRP
Anticorps anti MRR-eIF3b humain (Monoclonal de souris)	1/2000	Anticorps secondaire HRP
Anticorps anti MRR-eIF3g humain (Monoclonal de souris)	1/2000	Anticorps secondaire HRP
Anticorps anti eIF3j humain (TEBU)	1/2000	Anticorps secondaire HRP

Les membranes de nitrocellulose, après transfert ou fixation des protéines par dot blot, sont saturées par un tampon TNTlow (20mM Tris-HCl pH 8; 150mM NaCl; 5% Tween20) en présence de 5% de lait écrémé pendant 1h00 à température ambiante. Les anticorps sont dilués dans le même tampon et incubés à 4°C pendant un minimum de 4h. Les membranes sont lavées dans le tampon TNTlow sans lait, puis dans le TNTHight (TNTlow + 10% de tween20) avant d'être révélée par le réactif Luminex® (Invitrogen) substrat produisant un composé brun après action de l'enzyme HRP porté par l'anticorps. Afin d'augmenter la résolution il est possible d'utiliser un substrat ECL qui produit un composé fluorescent, révélé à l'aide d'un film photosensible (Ge Healthcare) en chambre noire.

I-3 Méthodes de caractérisation

I-3-A Spectroscopie d'absorbance

La concentration des solutions de protéine est déterminée par mesure du spectre d'absorbance (220-350nm). En utilisant le coefficient d'extinction molaire $DO = \epsilon \cdot l \cdot c$ des résidus Trp, Tyr et Cys à 280 nm (Pace *et al.* 1995), il est possible de calculer le coefficient pour chaque protéines d'après l'équation :

$$\epsilon_{280\text{nm}} (\text{M}^{-1} \text{cm}^{-1}) = N_{\text{Trp}} \times \epsilon_{\text{Trp}280 \text{ nm}} + N_{\text{Tyr}} \times \epsilon_{\text{Tyr}280 \text{ nm}} + N_{\text{Cys}} \times \epsilon_{\text{Cys}280 \text{ nm}}$$

avec $\epsilon_{\text{Trp}280 \text{ nm}} = 5\,500 \text{ M}^{-1} \text{cm}^{-1}$; $\epsilon_{\text{Tyr}280 \text{ nm}} = 1\,490 \text{ M}^{-1} \text{cm}^{-1}$; $\epsilon_{\text{Cys}280 \text{ nm}} = 125 \text{ M}^{-1} \text{cm}^{-1}$, et N_{Trp} , N_{Tyr} , N_{Cys} respectivement le nombre d'acides aminés Trp, Tyr et Cys de la protéine.

Protéines	MRR-eIF3b	MRR-eIF3g	eIF3b	eIF3j
$\epsilon_{280\text{nm}}$ (M-1 cm-1)	7450	8480	144185	28545

I-3-B Dichroïsme circulaire

Les spectres de dichroïsme circulaire (CD) ont été mesurés à l'aide d'un spectropolarimètre JASCO J-810. L'échantillon est contenu dans une cuve en quartz de 1 mm de trajet optique. Tous les échantillons ont été dialysés dans un tampon 20 mM Tris-HCl pH 7,5; 150 mM NaF; avant d'être analysés. Tous les échantillons protéiques sont ajustés à une concentration de 20 μ M sauf la protéine eIF3b qui est à 8 μ M.

I-3-C Spectroscopie de fluorescence des résidus tryptophanes

Les spectres de fluorescence ont été mesurés à l'aide d'un spectrofluorimètre Hitachi F2500 en utilisant une cuve en quartz de 1 cm. La longueur d'onde d'excitation est de 295 nm, le spectre d'émission a été analysé de 300 à 400 nm. Tous les échantillons ont été dialysés dans le même tampon précédemment utilisé pour les analyses de CD. Les concentrations de protéines utilisées sont 20 μ M pour le MRR-eIF3b et 8 μ M pour la protéine eIF3b entière.

I-3-D SEC-MALLS-RI

SEC-MALLS : Diffusion de lumière multi-angles couplé à la chromatographie d'exclusion de taille et à la réfractométrie. La chromatographie d'exclusion de taille est réalisée à l'aide d'une HPLC (chromatographie en phase liquide haute performance) avec une colonne de filtration sur gel superdex S200 (GE Healthcare). La diffusion de lumière statique est mesurée à 18 angles différents simultanément à l'aide d'un détecteur DAWN-EOS (Wyatt Technology Corp., Santa Barbara, CA) en utilisant un laser qui émet à 690 nm, l'excès d'indice de réfraction est mesurée à l'aide d'un réfractomètre différentiel RI2000b (Schambeck). Les données obtenues sont analysées avec le logiciel ASTRA V (WyattTechnology Corp.).

Echantillons	MRR-eIF3b	eIF3b	ARN
Concentration	10mg/ml	5mg/ml	2,5mg/ml

Tampon pour les protéines : 20mM Tris HCl pH 7,5; 150mM NaCl

Tampon pour les ARN : 20mM Tris-HCl pH 7,5; 100mM AcK; 200mM KCl, 2,5mM MgCl₂ (Kieft *et al.* 2001).

I-3-E Dynamic light scattering (DLS)

La technique de diffusion dynamique de la lumière (DLS), a été réalisée sur le nanoDLS (Malvern) et les échantillons sont préparés comme cela est décrit dans le paragraphe ci-dessus.

II Etude de l'IRES du VHC

II-1 Clonage des différents fragments d'IRES

Pour la plupart des fragments d'IRES, la séquence cible a été insérée dans un vecteur pUC18. La séquence cible est amplifiée par PCR qui permet l'insertion du site de restriction EcoRI précédé de la séquence promotrice de la T7 RNA polymérase en 3' et l'insertion du site de restriction HindIII en 5' de la séquence. Pour certains domaines, de petite taille, ou ne pouvant être amplifiés par PCR, j'ai utilisé la technique de gènes synthétiques afin de créer artificiellement la séquence d'ADN correspondant à la construction d'ARN souhaitée. Pour cela deux étapes de PCR sont nécessaires.

1er PCR : Génération du brin matrice.

Les oligonucléotides ont été synthétisés de telle sorte qu'ils se chevauchent et puissent former la séquence d'ADN matrice.

1 μ g de chaque oligo est utilisé pour réaliser la première PCR.

Nom	Volume
Oligonucléotides	1 μ l 1 μ g/ μ l
Master mix 2x Phusion GC rich	10 μ l
H2O	qsp 20 μ l

Cycle 1 (x30)	
96°C	2 min
72°C	1min

Après cette première étape, le mélange réactionnel est composé de plusieurs molécules d'ADN constituées par la combinaison des différents oligonucléotides. Afin d'amplifier la séquence d'intérêt, une seconde PCR est réalisée en utilisant les oligonucléotides spécifiques de chaque extrémité.

2nd PCR : Amplification de la cible

1 μ l de la première PCR est utilisé pour initier cette nouvelle PCR dans les conditions standard.

Cycle 1 (x3)		Cycle 2 (x30)		Cycle 3 (x1)	
96°C	2 min	96°C	2 min	72°C	10 min
72°C	1min	63-68°C	1min + 1s par cycle	15°C	n.a
		72°C	30s par kb + 1 s par cycle		

Cette seconde PCR permet d'amplifier la séquence cible et de la cloner par une approche clas-

sique dans le vecteur pUC18 comme cela a été décrit pour la constructions d'IRES classique. La liste des oligonucléotides est présentée dans l' **Annexe II** (page 184).

II-2 Biochimie des ARN

Tous les échantillons nécessaires à la production de l'ARN ont été traités dans les conditions RNase «free» : H₂O qualité RNA, filtration et autoclavage de tous les produits en solution, utilisation de plastiques stériles et à usage unique, cône à filtre pour les pipettes. De plus, toutes les étapes sont effectuées sous une hotte à flux laminaire dédiée à la production d'ARN décontaminé au RNAzap (SIGMA) après chaque utilisation.

II-2-A Préparation de la matrice ADN

Après purification de la matrice ADN, par utilisation d'un kit standard d'extraction (Qiagen ou Macheray Nalgen), l'ADN est linéarisé par action de l'endonucléase HindIII (NEB) à raison de 1 unité pour 100ng d'ADN dans le tampon de réaction adéquate (NEB2) pendant 12h à 37°C dans un volume de 200µl. Le produit de réaction est ensuite traité au phénol chloroforme saturé à l'alcool isoamylique volume à volume pendant plusieurs minutes afin d'éliminer la totalité de l'enzyme. La phase supérieure contenant l'ADN est ensuite précipité par ajout de 2,5 volumes d'éthanol 100% en présence de 300mM d'acétate de sodium à pH 5,3. Après une incubation à -20°C, l'ADN est centrifugé à haute vitesse à 4°C, le culot est lavé plusieurs fois à l'éthanol 70% avant d'être séché sous hotte RNase free.

II-2-B Protocole de transcription *in vitro*

Afin de réaliser la production d'ARN dans les meilleurs conditions possibles, tous les produits sont préparé à l'aide d'eau autoclavé RNase free. Tous les produits commerciaux sont également adapté à la production d'ARN. La transcription (Tableau IX) est réalisée sous une hotte dédié à la production d'ARN.

Tableau X : Protocole de la transcription *in vitro*

Tampon	Petite échelle	Grande échelle
1M Tris, pH 8,0	4 µl	400 µl
1M MgCl ₂	2,5 µl	250 µl
0,1M DTT (SIGMA)	0,5 µl	500 µl
0.1M Spermidine (SIGMA)	0,2 µl	20 µl
BSA 10mg/ml	1 µl	100 µl
0,2M NTP (Fermentas)	2 µl	200 µl
0,01% Triton	10 µl	1 ml
40% PEG 6000	20 µl	2 ml
EMBL T7 Polymerase (10mg/ml)	10 µl	100 µl
ADN linéarisé et purifié	10 µg	1 mg
RNAguard (GE Healthcare)	5 µl	15 µl
H ₂ O	qsp 100 µl	qsp 10ml
Temps d'incubation	3h à 37°C	12h à 37°C sous agitation

II-2-C Obtention d'ARN radioactif

Afin d'obtenir de l'ARN uniformément marqué au $\alpha^{32}\text{P}$, nous utilisons un nucléotide marqué : le $\text{UT}\alpha^{32}\text{P}$ (3000Ci/mmol) (Perkin Elmer). Ce nucléotide est additionné au mélange réactionnel lors de la transcription in vitro.

Nucléotides ATP, CTP, GTP : 4mM final

Nucléotide UTP froid : 0,4mM final

Nucléotide UTP radioactif $\alpha^{32}\text{P}$: 3 μM final soit environ 3 μl

La réaction est effectuée dans un volume de 20 μl , en pièce radioactive.

II-3 Purification

II-3-A Petite quantité d'ARN

L'ARN est purifié sur gel d'acrylamide en conditions dénaturantes. L'électrophorèse est réalisée en TBE 1x (10.8g Tris base; 5.5g Boric acid; 4ml 0.5M EDTA pH 8,0) en présence de 8M d'urée finale avec 10% d'acrylamide/Bis-acrylamide 29/1. Le gel est chauffé 30 minutes à 1W/cm² avant d'être utilisé. Les échantillons sont mélangés à un tampon de dépôt contenant 1mM EDTA pH8 et 8M d'urée avant d'être chauffés à 95°C pendant 2 minutes. Les échantillons sont déposés sur le gel et migrent sous 60W.

Cas des ARN radioactifs : Après la migration, le gel est entouré d'un film protecteur avant d'être exposé plusieurs minutes en chambre noire contre un film photosensible (GE Healthcare digital X-Ray). Le film est révélé par des produits photographiques, nous permettant de localiser les bandes correspondant aux ARN amplifiés.

Cas des ARN non radioactif : Après la migration, le gel est entouré d'un film protecteur avant d'être déposé sur une plaque phosphorescente et illuminé au UV à 265nm. Cette technique permet de visualiser l'ARN par ombrage (UV Shadowing).

Les fragments d'ARN radioactifs sont excisés du gel et incubés dans un tampon de diffusion (AcNH_4 500mM; AcMg 10mM; SDS 0,1%; EDTA 0,1mM) pendant 1 nuit à température ambiante. L'ARN est ensuite précipité par ajout de 2,5 Volumes d'éthanol 100% et de 20 μg de glycogène. Après une incubation à -20°C, le mélange est centrifugé à 16000g à 4°C, le culot est lavé plusieurs fois à l'éthanol 70% avant d'être séché sous hotte à température ambiante.

II-3-B Grande quantité d'ARN

Le produit de transcription est additionné de 4M urée final en présence de 10mM d'EDTA avant d'être dénaturé pendant 10 minutes à 95°C. Après cette étape, l'échantillon est centrifugé à température ambiante pendant 20 minutes à 16000g puis injecté sur une colonne d'échange d'ions Q sépharose (GE Healthcare) (50ml pour 10ml de transcription) est équilibrée dans le tampon I (20mM Tris HCl pH7,5; 10mM EDTA; 4M d'urée, filtré, autoclavé).

La colonne est lavée par 2 volumes de colonne du même tampon, puis par 2 volumes de colonne du tampon I additionné de 500mM de NaCl. L'ARN est ensuite élué par gradient linéaire de NaCl (500mM à 750mM) sur 50 volumes colonne. La collecte des fractions est réalisée par pallier de 5ml.

Les fractions sont analysées sur gel d'acrylamide/bisacrylamide 29/1 10% en condition dénaturante (8M d'urée). Les fractions d'intérêt sont rassemblées en dialysées contre 1L d'H₂O avant d'être partiellement concentrées par lyophilisation. L'ARN est ensuite resuspendu dans 5ml d'H₂O à 70°C avant d'être injecté sur une colonne de chromatographie d'exclusion de taille S200 préparative de 120ml (GE Healthacare) préalablement équilibrée en acétate de sodium 300mM, 10mM EDTA. Les fractions d'intérêts sont analysées sur gel d'acrylamide en condition dénaturante avant d'être rassemblées en précipitées par addition de 0,7 volume d'isopropanol 100%. Après centrifugation à 30000g, le culot obtenu est lavé 2 fois à l'éthanol 70% avant d'être séché à l'air libre sous la hotte RNase free. Les ARN sont contrôlés sur gel d'acrylamide 10% en condition native (0,1W/cm²) et/ou dénaturante (8M d'urée). Les échantillons sont fractionnés en aliquot de 1mg, avant d'être une dernière fois lyophilisés, pour être conservés déshydratés à -80°C.

III Etudes d'interaction de complexe RNP

III-1 Activation de l'ARN

L'ARN est repris dans la moitié du volume final souhaité par de l'H₂O traité au DEPC à 50°C. Après 10 minutes, le tube est agité doucement pendant 2 minutes avant d'être chauffé à 95°C pendant 1 minute. L'ARN est ensuite refroidi brutalement par ajout de tampon de interaction/repliement froid Tris-HCl (pH 7,5) 100mM AcK; 200mM KCl; 2,5mM MgCl₂ et 1mM DTT. L'ARN ainsi activé est incubé 10 minutes avant utilisation et doit être conservé à température ambiante. L'échantillon est ensuite centrifugé à température ambiante pendant 10 minutes à 16000 g avant utilisation.

III-2 Filter Binding Assay

Les différents échantillons de protéine sont dilués en série (de 10⁻⁴M à 10⁻⁹ M) dans un volume final de 180 µl de tampon d'interaction suivi par l'addition de 20 µl d'ARN radioactif activé (minimum 5000 cpm). Une incubation de 1 h à température ambiante est réalisée dans le tampon de binding. Le mélange réactionnel est ensuite filtré à travers un filtre de nitrocellulose pré-imbibé dans le tampon d'interaction (Nitrocellulose Whatman, pores de 0.45 µm, 25 mm de diamètre). Seul l'ARN complexé à l'échantillon protéique est retenu sur le filtre, qui peut être détecté par la méthode Cerenkov dans un compteur à scintillation.

III-3 Retard sur gel «EMSA»

Les différents échantillons de protéine sont dilués en série (de 10^{-4} M à 10^{-9} M) dans un volume final de 20 μ l de tampon de binding suivi par l'addition de 2 μ l d'ARN radioactif activé (minimum 5000cpm). Une incubation de 1 h à température ambiante est réalisée dans le tampon de binding. Après incubation, les échantillons sont additionnés de du bleu de chargement ARN natif (Sucrose 6%, bleu de méthylène et de xylène cyanol). Les complexes sont déposés sur un gel d'acrylamide/bisacrylamide 29/1» de 6 à 12% en fonction de la résolution souhaitée. La migration est effectuée en tampon TA (10 mM Tris-acétate pH 7,8; 66 mM AcK; 10 mM AcMg) pendant 4h à 100V (0,1W/cm²) en chambre froide (température du gel environ 20°C). Le gel est ensuite fixé par une solution contenant 10% acide acétique et 10% éthanol pendant 10 minutes avant d'être séché 1h00 sur un sécheur de gel sur du papier Whatman 3MM. Le gel sec est ensuite exposé à un film photographique dans une cassette pendant quelques heures avant d'être révélé.

III-4 Etude RMN

La protéine MRR-eIF3b est utilisée à 1.10^{-4} M en tampon de binding et l'ARN est concentré à 15mg/ml soit $1,5.10^{-4}$ M. Le matériel biologique dont nous disposons nous a permis de réaliser 4 points de titration. L'ARN déshydraté est additionné de 200 μ l d'H₂O chauffé 3 minutes à 95°C avant d'être refroidi par l'addition de 200 μ l de tampon de binding froid.

IV Small angle X-Ray Scattering SAXS

IV-1 Protéine en solution

Les échantillons sont décongelés extemporanément et dessalés sur colonne G25 (GE Healthcare) par centrifugation avant utilisation. La concentration de chaque échantillon est mesurée par spectrophotométrie d'absorbance sur un spectrophomètre nanogoutte.

Nom	Concentration	Tampon	Nombre de mesure
eIF3b	5mg/ml, 2,5mg/ml	HEPES 20mM pH 7,5; 150 mM NaCl	10
MRR-eIF3b	10mg/ml, 5mg/ml	20mM Tris HCl pH 7,5; 150mM NaCl	5
BSA	3,8mg/ml	20mM Tris HCl pH 7,5; 150mM NaCl	5

IV-2 ARN en solution

Les échantillons d'ARN sont préparés extemporanément comme cela est décrit dans le protocole d'activation de l'ARN.

Nom	Concentration	Tampon	Nombre de mesure
ARN	10mg/ml, 5mg/ml, 2,5mg/ml	Tris-HCl (pH 7.5) 100mM AcK, 200mM KCl; 2,5mM MgCl ₂	2
ARN	10mg/ml, 5mg/ml, 2,5mg/ml	H2O	2

IV-3 Complexe Ribonucléoprotéique en solution

Les échantillons sont préparés extemporanément et incubés 30 minutes à température ambiante avant utilisation. Chaque fragment d'ARN a été utilisé pour réaliser un complexe avec chacune des 3 protéines : eIF3, eIF3b et MRR-eIF3b

Nom	Concentration de protéine	Concentration d'ARN	Rapport molaire Protéine / ARN
IRES / eIF3	1.10-5M	1.10-4M	1 : 10
IRES / eIF3b	2.10-5M	1.10-4M	1 : 5
IRES / MMM-eIF3b	5.10-4M	1.10-4M	1 : 0,5

RESULTATS & DISCUSSIONS

Etude structurale par SAXS de
l'IRES du VHC

Introduction

L'IRES du VHC est localisé en 5' du génome et comprend 332 nucléotides agencés en quatre domaines notés de I à IV. Certains fragments ont été étudiés de manière individuelle et leurs structures ont été résolues par utilisation des techniques de cristallographie et/ou de RMN. Néanmoins, les informations concernant l'IRES VHC - seul en solution - ne sont pas disponibles actuellement. Ces données permettraient de faire la lumière sur les mécanismes dynamiques conduisant à la formation des structures nécessaires au recrutement des différents facteurs (Kieft *et al.* 2002). Une étude réalisée en 2001 étudiant l'IRES du VHC par microscopie électronique, a permis d'identifier 3 domaines (II, III, IV). Les auteurs ont ainsi pu proposer un modèle en 2D de l'IRES du VHC qui est toujours le modèle de référence (**FIG20**). La présentation de chaque domaine a fait l'objet d'un paragraphe dans l'introduction et indique la fonction de chaque domaine et sous domaine dans les phases de l'initiation de la traduction. Cependant, les informations structurales de l'IRES du VHC ne sont que fragmentaires, et peu d'équipes tentent de comprendre cette structure sous sa forme entière. Afin de mieux appréhender la dynamique moléculaire de cet ARN, nous avons utilisé une approche structurale originale afin d'étudier l'IRES du VHC en solution. Nous avons utilisé la technique de diffusion aux rayons X aux petits angles (SAXS : Small Angle X-rays Scattering). Cette technique a déjà été utilisée

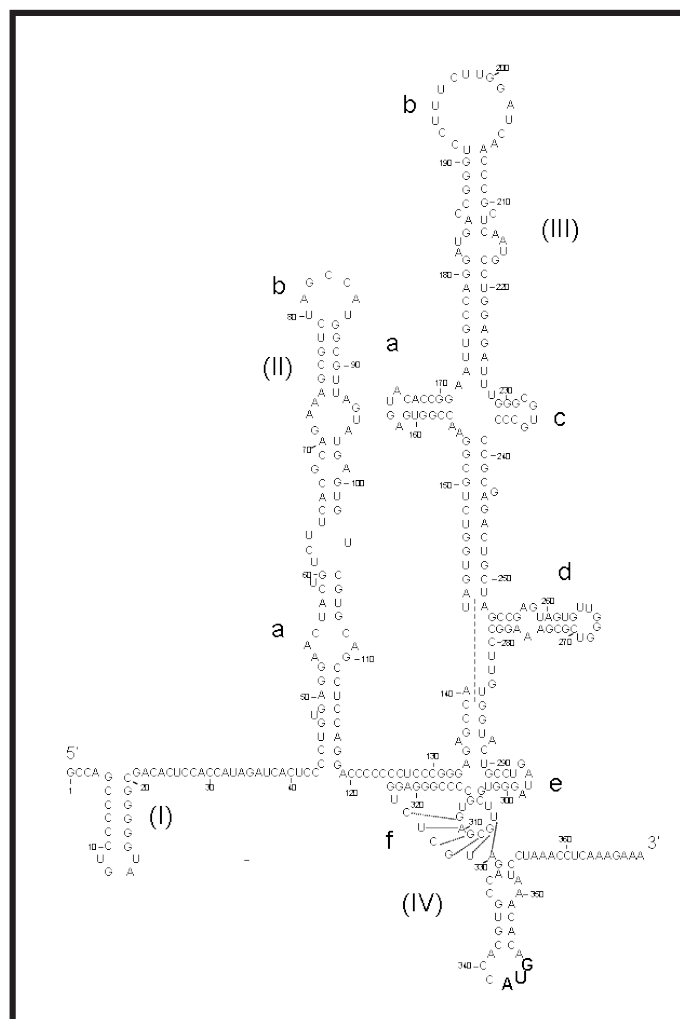


FIG20 : Présentation du modèle en 2D de l'IRES du VHC.

pour étudier la forme entière de l'IRES en solution (Kieft *et al.* 1999). Cependant les auteurs ne sont pas arrivés à obtenir des échantillons de qualité suffisante pour réaliser des mesures de diffusion permettant la modélisation de cet ARN. Depuis quelques années, la technique de SAXS a fait de gros progrès au niveau technique et informatique dans le traitement des données et la modélisation. Plusieurs enveloppes d'acides nucléiques ont été étudiées par cette technique (Putnam *et al.* 2007). Cependant les structures étudiées étaient des molécules de petite taille et très structurées à l'image du ribozyme Tetrahymena (Russell *et al.* 2000). Ce qui n'est pas le cas de l'IRES du VHC qui est une structure de grande taille prédite non structurée au niveau du domaine IV par exemple. Afin d'étudier la faisabilité de cette technique sur l'étude structurale de l'IRES du VHC en solution et de tester la pureté de nos échantillons, nous avons réalisé une première série de mesure sur la ligne ID2 de l'ESRF (European Synchrotron Radiation Facility, Grenoble). A la suite de cette expérience, les paramètres biophysiques de l'ARN ont pu être mesurés. Ceux-ci indiquent la présence d'une molécule allongée (250Å) et d'un Rg supérieur à 70Å. Les résultats obtenus étant prometteurs, nous avons décidé d'utiliser cette technique pour réaliser une étude complète de l'IRES en solution. Afin de réaliser une étude complète de cet ARN viral, une demande à l'ESRF nous a été octroyée avec 24h de temps de faisceau. Ces 24h nous ont permis de réaliser l'acquisition de 150 courbes de diffusion des différents échantillons (ARN, protéines et complexes). Cet ARN étant de grande taille nous avons opté pour une étude séquentielle des différents fragments. Afin d'étudier l'IRES en détail ainsi que la dynamique de certains domaines, j'ai découpé l'IRES en sept fragments en me basant sur le modèle de l'IRES en 2D (**Tableau XI**). Cette découpe moléculaire nous a permis d'étudier séparément les domaines II et III ainsi que la combinaison d'autres domaines (**FIG21**). Chacun de ces fragments a été utilisé pour mesurer une courbe de diffusion par la technique de SAXS. Les expériences de diffusion de SAXS ont été réalisées sur la ligne de lumière ID14.3 de l'ESRF à une longueur d'onde de 0,995Å. Après obtention des courbes de diffusion de chaque fragment d'ARN, la qualité des résultats a été évaluée, avant de générer un modèle de l'IRES en solution.

Tableau XI : Liste des constructions réalisées pour l'étude de SAXS

Nom	Taille (nt)	MW (kDa)
IRES II	77	24
IRES IIIabc	132	44
IRES IIIabcd	170	56
IRES IIIdefV	143	47
IRES III	210	69
IRES III-IV	250	82
IRES II-III	274	90
IRES II-III-IV	330	110

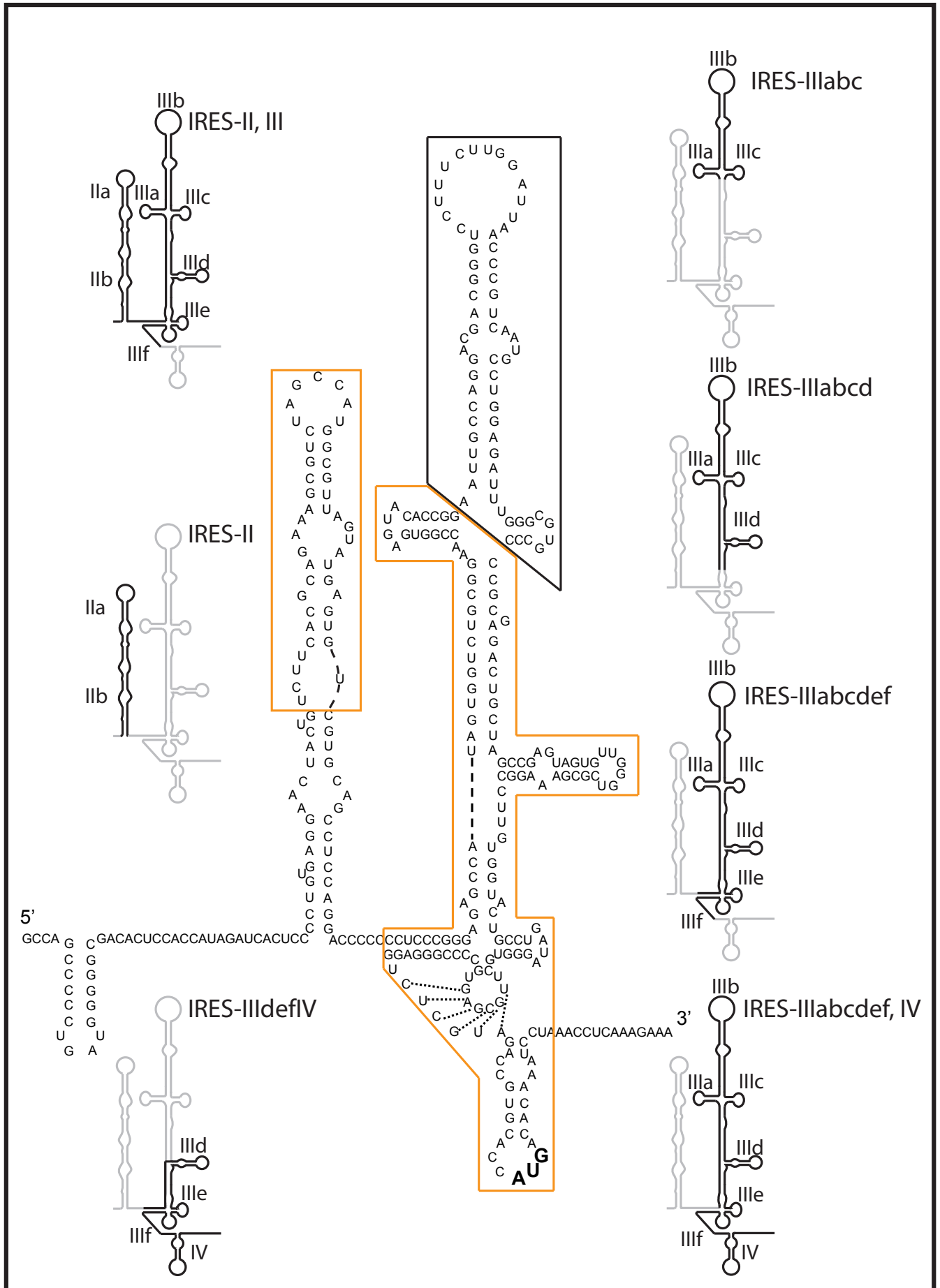


FIG21 : Dissection moléculaire de l'IRES du VHC.

En noir, région impliquée dans l'interaction du facteur eIF3.

En orange, régions impliquées dans l'interaction avec 40S.

D'après la revue de Lukavsky publiée en 2008 : Structure and function of HCV IRES domains.

I Dissection moléculaire de l'IRES et obtention des échantillons.

I-1 Clonage, production et purification

Les séquences correspondant à l'IRES, à ses domaines ou sous-domaines ont été clonées dans un vecteur pUC18 (**FIG22A**) en aval d'un promoteur T7. L'utilisation d'une enzyme de restriction (HindIII) permet de linéariser le vecteur en 3' de la séquence d'ARN correspondant. L'ARN polymérase T7 permet de produire efficacement l'ARN qui sera purifié par chromatographie d'échange d'ions en conditions dénaturantes (**FIG22B, C, D**).

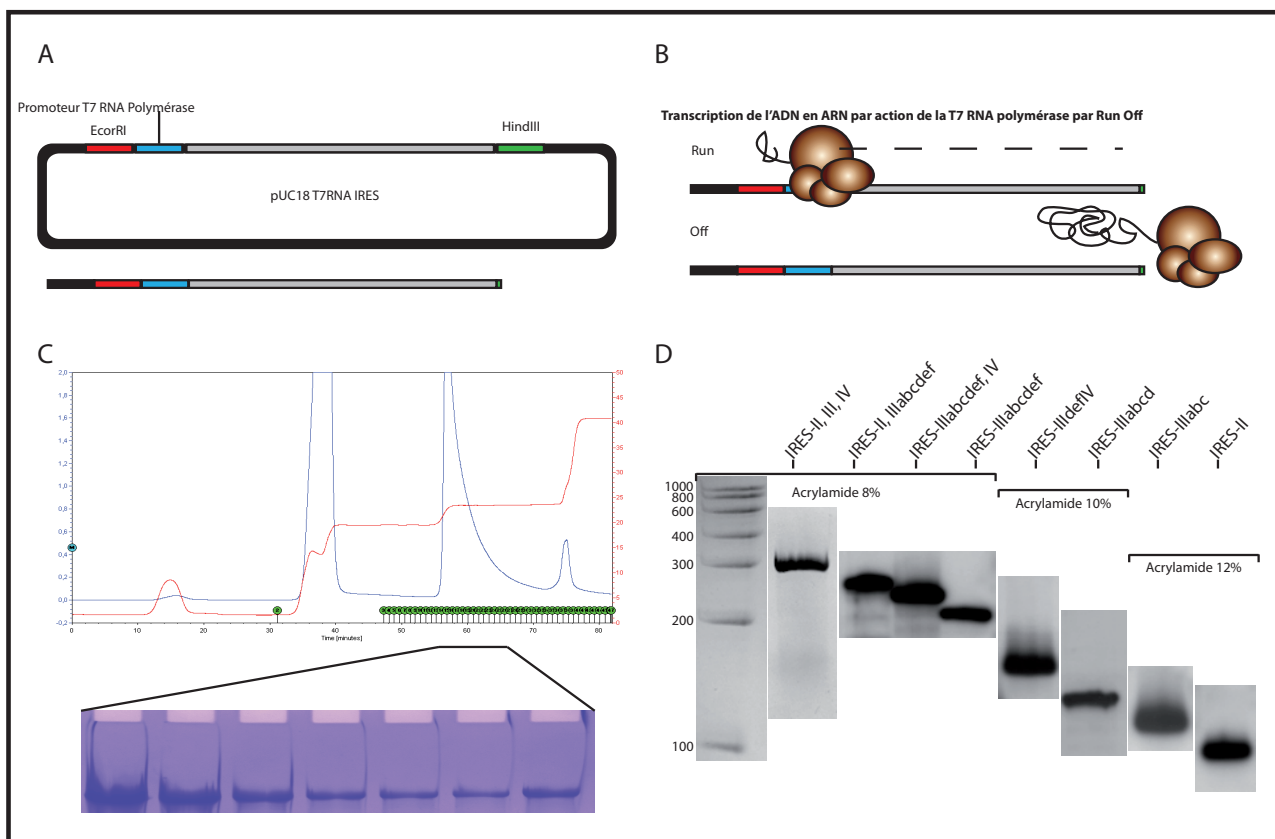


FIG22 : Clonage, transcription et purification des différents domaines de l'IRES du VHC.

A : Construction du vecteur puC18 et insertion des séquences d'ARN.

B : Linéarisation de l'ADN et transcription in vitro par un run off.

C : Purification par chromatographie d'échanges d'ions en condition dénaturante et gel analytique des fractions purifiées en gel dénaturant.

D : Gel analytique des ARN purifiés par électrophorèse en condition dénaturante.

I-2 Caractérisation

Avant d'utiliser ces échantillons pour la collecte de données de diffusion, nous avons étudié l'homogénéité de chaque construction d'ARN par utilisation de la diffusion dynamique de lumière (DLS : Dynamic Light Scattering). Il s'agit d'une technique biophysique rapide qui permet de mesurer la taille apparente d'une particule. Cette mesure, utilisée en routine, m'a permis de m'assurer de la monodispersité de chacun de mes échantillons (**FIG23**). Les mesures de chaque construction ont été réalisées dans le tampon de refolding (Mat&met) contenant du magnésium. On peut noter que le domaine II présente le plus petit diamètre tandis que la présence du domaine IV semble au contraire augmenter fortement le diamètre de l'ARN. Cependant, la technique de DLS n'est pas très adaptée à l'étude des nucléoides. Les informations qui seront retenues ne seront que qualitatives et nous indiquent que chaque échantillon est monodisperse (minimum 95%).

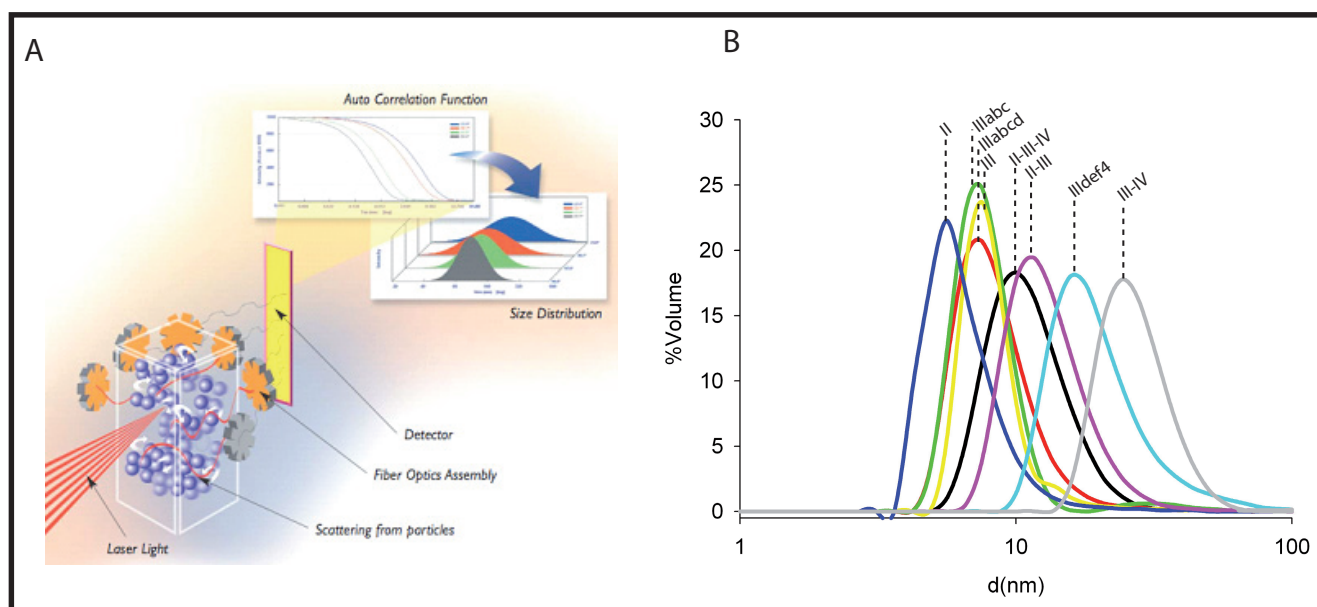


FIG23 : Résultats de Dynamic Light Scattering et vérification de l'homogénéité des échantillons d'ARN.

A : Principe de la DLS : Les échantillons sont introduits dans la cuve d'analyse à température ambiante. Les particules sont illuminées par un laser. L'intensité de la lumière diffusée dépend de la taille de la particule et de la viscosité de milieu. La mesure de la fluctuation de l'intensité couplée au mouvement Brownien permet de mesurer la taille de la particule. Cette technique rapide permet d'apprécier la taille d'une molécule, de connaître son rayon hydrodynamique et d'estimer l'homogénéité de l'échantillon analysé.

B : Chaque construction d'ARN est analysée par DLS à 20°C dans le tampon de binding. Les résultats indiquent que tous les échantillons sont homogènes et monodisperses à plus de 95%

II La technique de SAXS

II-1 Principe du SAXS

La diffusion aux petits angles (SAS : Small Angle Scattering) est une méthode d'analyse des macromolécules biologiques qui permet d'obtenir des informations sur la taille et la forme de ces antitées en solution (**Tableau XII**).

Tableau XII : Liste des différents paramètres biophysiques obtenus par SAXS

Paramètres	Résultats attendus et interprétations
Expérimentaux	
Guinier plot	L'obtention d'une courbe non linéaire indique la présence d'aggrégation.
R _g	Le rayon de gyration, qui est calculé à partir du graphique de Guinier, est représentatif de la taille de la molécule.
I(0)	Cette constante est proportionnelle à la taille de la molécule et permet de connaître la masse moléculaire et d'estimer la concentration de l'échantillon.
D _{max}	Représente la taille maximum de la molécule qui doit être corrélée par une distribution de fonction homogène.
P(r)	Représente la distribution de fonction de la courbe de diffusion qui doit être homogène.
Modélisation	
(R or χ^2)	Valide la qualité du modèle
NSD	Permet de visualiser la convergence des modèles construits, et de sélectionner des sous groupes de modèles, indiquant la flexibilité de la molécule par exemple.
R _g modèle	Permet de vérifier le R _g de la protéine par rapport à la théorie
D _{max} modèle	Permet de vérifier le D _{max} par rapport à la théorie
P(r) modèle	Permet de valider les modèles atomiques

La **figure 24** présente une expérience type : un faisceau monochromatique (neutrons ou rayons X) illumine un échantillon qui diffuse élastiquement la radiation incidente. L'intensité diffusée aux petits angles est mesurée sur un détecteur qui transforme les données en une courbe de diffusion. Cette courbe (intensité diffusée en fonction de l'angle de diffusion) est utilisée pour créer des modèles à basse résolution des molécules analysées. A la suite de l'obtention des données brutes de diffusion, il est nécessaire d'utiliser plusieurs outils informatiques qui permettront de traiter les données.

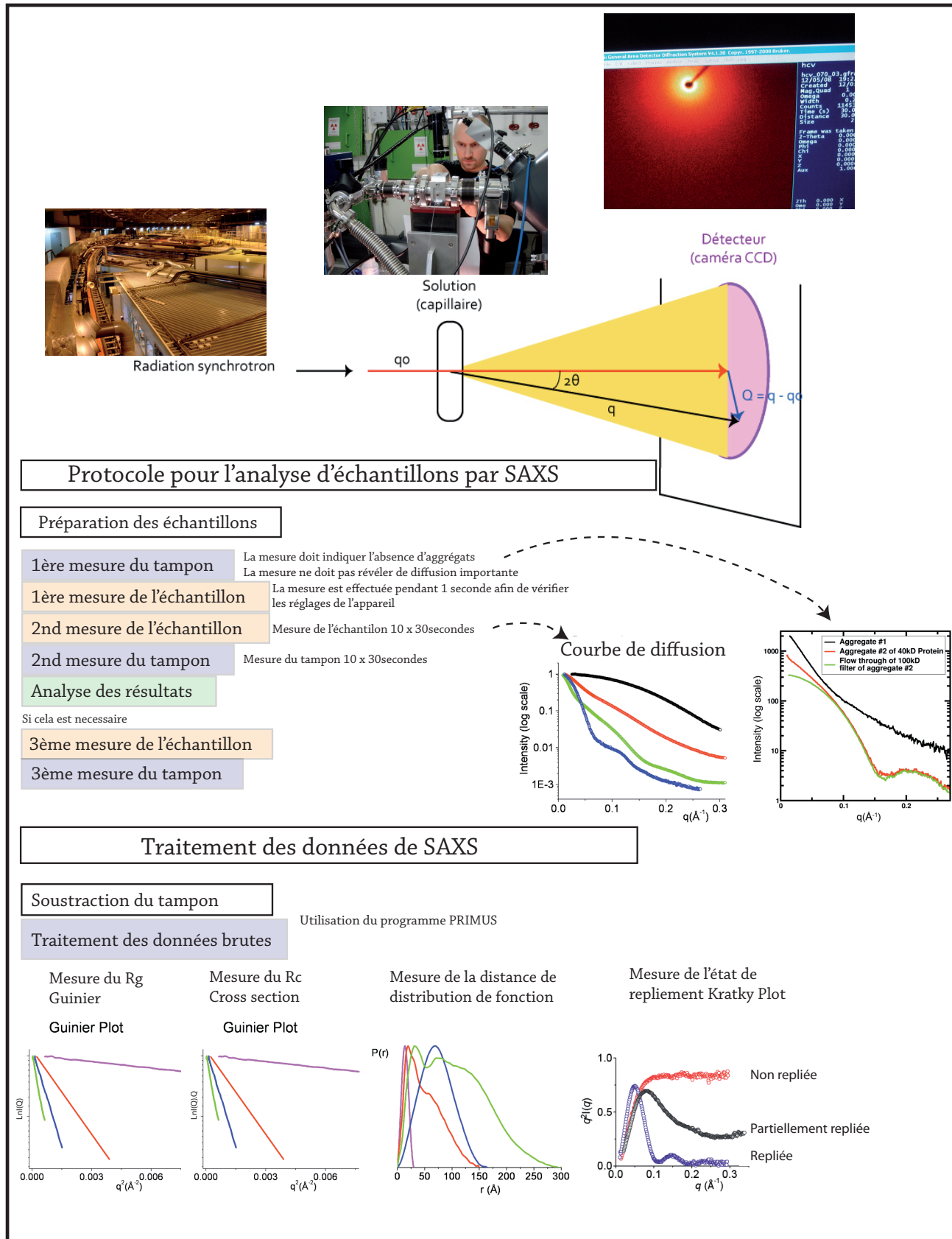


FIG24 : Principe du SAXS, exemple de la ligne ID14.3 de l'ESRF de Grenoble.

Un faisceau monochromatique illumine un échantillon présent dans un capillaire. Celui-ci, après excitation, diffuse élastiquement. L'intensité diffusée aux petits angles est alors analysée et intégrée par un détecteur pour donner une courbe de diffusion. Les courbes ainsi obtenues nous permettent d'obtenir de nombreuses informations sur la taille, la forme, l'état de repliement ou la présence d'agrégats. (Koch *et al.* 2003; Putnam *et al.* 2007)

II-2 Les différents outils informatiques

L'utilisation de chaque programme nécessite la préparation des données de manière rigoureuse. Comme cela est présenté dans le tableau ci-dessous, chaque programme (**Tableau XIII**) est dédié à une application particulière, qui permet de traiter les données brutes ou de générer des modèles à basse résolution *Ab initio* par exemple.

Tableau XIII : Fonction des différents programmes

Nom	Fonction
PRIMUS	Mise à l'échelle des courbes. Détermination du Rg à partir de l'approximation de Guinier.
GNOM	Calcul de la P(r), du Rg et du Dmax.
Sasplot	Permet de visualiser toutes les courbes générées par les programmes
DAMMIN, DAMMIF, GASBOR	Modélisation ab initio
CREDO	Modélisation d'un ou plusieurs domaines manquants dans une ou plusieurs structures cristallographiques
CRYSOL	Calcul de la courbe de diffusion théorique correspondant à une structure cristallographique
DAMAVER	Superposition des modèles deux à deux et détermination des valeurs du NSD (Normalized Spatial Discrepancy)
SUPCOMB	Superposition des modèles obtenus à partir de la courbe de diffusion avec une structure atomique.

Tous les programmes utilisés ont été générés par l'équipe de Svergun (<http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html>)

Pour une revue détaillée, lire (Koch et al. 2003).

II-3 Principe de la construction de modèles

La construction de modèle *Ab-initio* (traitement des données sans modifications) consiste à utiliser des programmes qui vont générer des modèles représentatifs des molécules en solution. Concrètement, les programmes comme DAMMIN (Svergun 1999) vont générer une sphère de taille déterminée par le Dmax de la molécule préalablement calculé pour chaque échantillon (à l'aide du programme GNOM). La sphère générée par le programme DAMMIN est composée de billes représentant des carbones alpha « fictifs ». Ainsi, le programme va enlever des billes de part et d'autre de la sphère de manière aléatoire en ajustant la courbe théorique à la courbe expérimentale. Il est donc nécessaire d'exécuter un grand nombre de fois ces programmes pour obtenir une bonne représentation de la molécule. A la fin de chaque cycle, le programme génère un fichier contenant une structure 3D représentant la forme de la molécule ainsi que plusieurs fichiers nous permettant de valider la précision du modèle. A la suite de la génération de ces nombreux modèles, il faut les comparer les uns aux autres pour établir une

bonne représentation statistique. Le programme DAMAVER va permettre de comparer tous les modèles entre eux, et de trouver une solution unique moyenne. Afin de discriminer les modèles ayant une solution trop éloignée du modèle moyen, nous utilisons le « Normalized Spatial Discrepancy » (NSD) (M. B. Kozin et D. I. Svergun) de chaque modèle. Ce chiffre correspond à l'écart mesuré entre chaque modélisation. D'après la théorie, les modèles ayant un NSD inférieur à 0.7 peuvent être assemblés en un seul modèle et correspondre à la solution la plus probable (Putnam *et al.* 2007). Les modèles possédant un NSD supérieur à 1, représente plusieurs autres solutions qui sont éloignées du modèle moyen choisi comme référence. Il est donc possible de séparer ces modèles et de relancer le programme DAMAVER qui va sélectionner un autre modèle de référence. Après plusieurs traitements, il est ainsi possible d'obtenir plusieurs modèles représentant tous la forme de la molécule en solution dans différentes configurations. Si les modèles sont trop différents les uns des autres, cela indique la présence d'un état dynamique de la molécule en solution. Les résultats qui sont présentés dans le paragraphe suivant sont le résultat de plusieurs cycles de modélisations réalisés à partir de 50 modèles individuels pour chaque fragments.

III Modélisation des domaines individualisés

III-1 Technique de modélisation

Afin de générer le modèle de chaque construction d'IRES en solution, nous avons utilisé le programme DAMMIN, couplé au programme DAMAVER. Pour chaque modèle, les différents paramètres biophysiques ont été mesurés. On peut constater que les tailles des domaines mesurés sont réparties entre 100 et 200Å avec des Rg compris entre 30 et 70 Å. Pour chaque modèle, il a été possible de mesurer le «I0» qui permet de calculer la masse et la concentration des échantillons mesurés (**FIG25**).

III-2 Présentation des modèles

- Le domaine II est composé de trois régions formant une légère courbure caractéristique. Il possède un Rg de 33Å et un Dmax de 97Å. Ce motif fera l'objet d'une étude dynamique dans le chapitre suivant (**FIG25**).
- Le domaine IIIabc est composé de deux régions séparées par une courte distance en formant une légère courbure. Il possède un Rg de 43 Å et un Dmax de 115 Å (**FIG25**).
- Le domaine IIIabcd est composé d'une région supplémentaire correspondant au domaine IIIId. On retrouve la forme du domaine IIIabc, qui semble être plus allongée et ne forme plus cette courbure. On peut noter que le Rg et le Dmax augmente en présence du domaine IIIId. Cette

construction contient en effet 48 nucléotides de plus que le domaine IIIabc. Ces données semblent compatibles avec la théorie et nous permettent de localiser le domaine IIIId (**FIG25**).

- Le domaine IIIdefIV forme un demi-cercle et se compose de trois régions. D'après les résultats, on peut identifier une région correspondant au domaine IIIId ce qui nous permet d'orienter la molécule et de localiser les domaines IIIe, IIIf et IV. Par ailleurs, on peut remarquer que les paramètres biophysique R_g et D_{max} sont identiques aux paramètres mesurés pour le domaine IIIabcd. Cela nous permet d'estimer la taille du domaine IIIabcdef qui doit posséder un D_{max} entre 150 et 200 Å. Cependant l'obtention de ce modèle ne permet pas de localiser avec précision le domaine IV (**FIG25**).

- Le domaine IIIabcdef nous permet de retrouver tous les sous domaines (abcdef). On note cependant que la forme du domaine IIIabcd n'est pas retrouvée en présence des domaines IIIef. De plus, la molécule semble plus compactée dans la région IIIabcd, nous indiquant la présence d'une influence des domaines IIIef sur la structuration du domaine III. Les paramètres biophysique confirment la présence d'une molécule plus allongée que le domaine IIIabcd avec un R_g de 50 Å et un D_{max} de 180 Å (**FIG25**).

- Le domaine III-IV nous permet de localiser le domaine IV par rapport au domaine III. Celui est en effet éloigné du domaine IIIabc et se situe après le domaine IIIf. Cependant, la localisation précise du domaine IV est plus difficile, même si l'on retrouve tous les sous domaines du domaine III. De plus, nous pouvons noter que le niveau de compaction du domaine III semble augmenter en présence du domaine IV (**FIG25**).

- Le domaine II-III nous permet de localiser le domaine II par rapport à tous les sous-domaines III. L'analyse du modèle indique que le domaine II est bien localisé au niveau des domaine IIIe et IIIf. De plus, nous pouvons noter que R_g et le D_{max} sont identiques à ceux du domaine III-IV (**FIG25**). Ces résultats indiquent que le domaine II n'allonge pas la molécule dans ces conditions.

D'après les courbes de diffusion obtenues, j'ai été en mesure de calculer la masse moléculaire de chaque fragment d'ARN en utilisant la mesure du I0. La mesure de la masse moléculaire par le I0 est directement proportionnelle à la concentration des échantillons. La concentration de chaque échantillon a été déterminée par spectroscopie d'absorbance à 260 nm à l'aide d'un spectrophotomètre NanoGoutte. Le résultat présentant la masse moléculaire de chaque fragment est reporté dans le tableau de la figure 25. Les mesures effectuées indiquent qu'il est possible d'obtenir la masse de chaque ARN avec une erreur inférieure à 10% pour les fragments de grande taille. Ces résultats sont représentatifs de la qualité des mesures et indique la ro-

bustesse de l'expérience. D'après l'analyse de chaque construction, il est possible de localiser les domaines II, III et IV ainsi que tous les sous-domaines du II et du III. Concernant le domaine IV, celui-ci semble être situé proche du domaine IIIf et présente une structure peu compacte. La présence des domaines II et IV semble influencer sur le niveau de compaction et sur l'arrangement du domaine III. L'analyse des différents domaines et combinaison des domaines présentée ici, nous permettent d'imaginer la forme de l'IRES en solution. Afin de valider ces résultats, nous nous sommes basés sur les données à haute résolution du domaine II, pour les confronter aux informations biophysiques mesurées par SAXS.

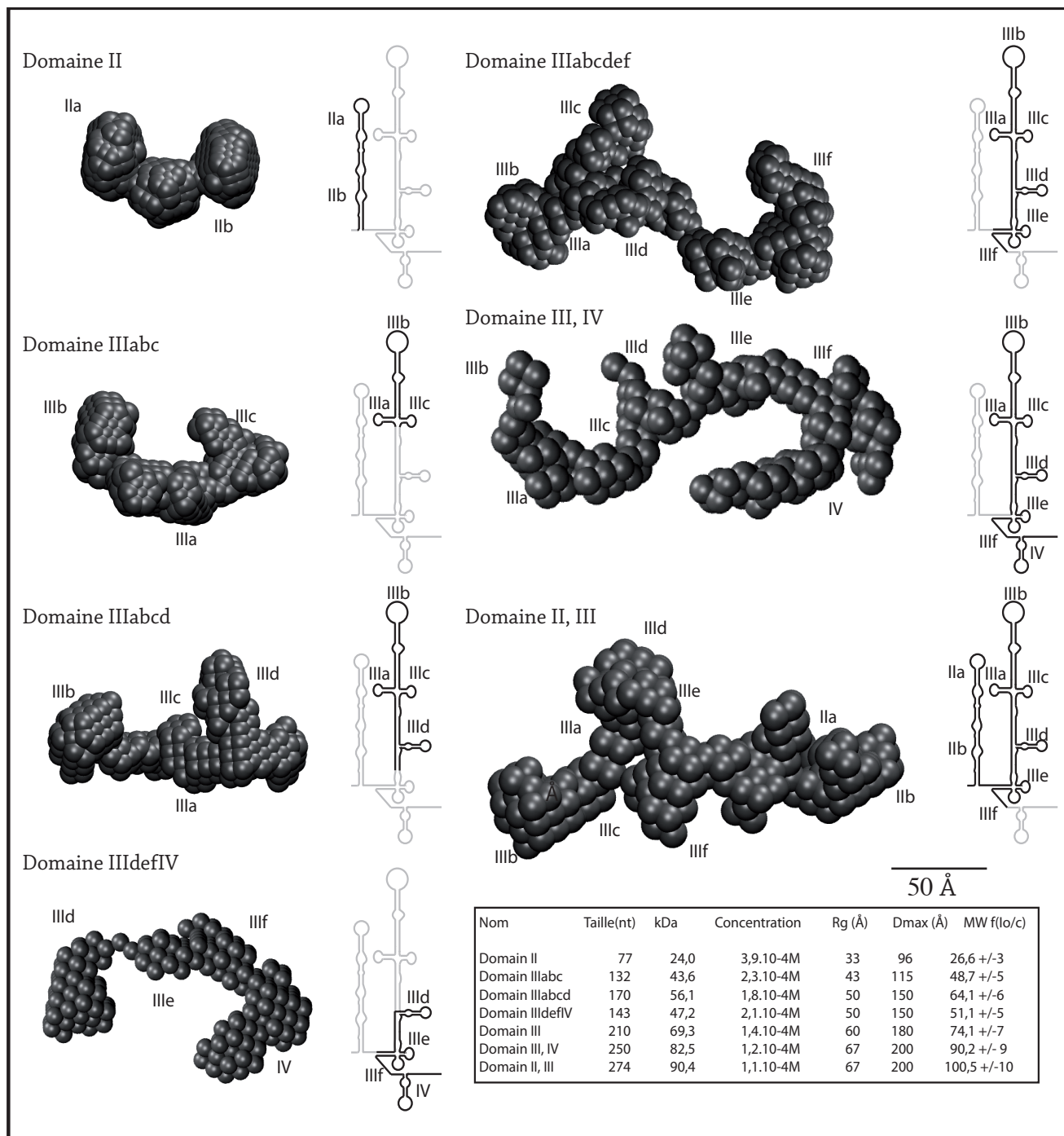


FIG25 : Dissection moléculaire de l'IRES du VHC et présentation des propriétés biophysiques. Chaque construction a été modélisée à l'aide du programme DAMMIN-53 (Svergun). Pour chaque domaine, 50 modèles ont été réalisés de manière indépendante. Tous les modèles ont ensuite été comparés les uns aux autres à l'aide du programme DAMVER (Svergun). Les modèles présentés sont représentatifs des modèles ayant un NSD inférieur à 0.7 (meilleure solution). Tous les modèles sont orientés en fonction des domaines II et III.

IV SAXS versus RMN : étude du domaine II

Les résultats présentent l'enveloppe et la structure à haute résolution du domaine II et sont reportés dans la figure **FIG26A**. On peut noter la présence d'une forme caractéristique du domaine II en faucille qui est retrouvée dans les modèles de SAXS ou de RMN. La résolution obtenue par la technique de SAXS est faible et ne permet pas de d'apprécier des détails inférieurs à 15Å (**FIG26B**). Cependant on peut remarquer que les modèles se superposent et présentent une courbure identique (**FIG26A et B**). Afin d'étudier et de comparer plus précisément ces résultats, j'ai effectué une mesure de la diffusion du domaine II dans les mêmes conditions que celles utilisées pour la RMN (Structure en rouge FIG 26B). Pour comparer les données de RMN et de SAXS j'ai utilisé le programme CRY SOL. Ce programme permet de générer une courbe de diffusion théorique (**FIG26C**) à partir d'une structure à haute résolution (**FIG26A**).

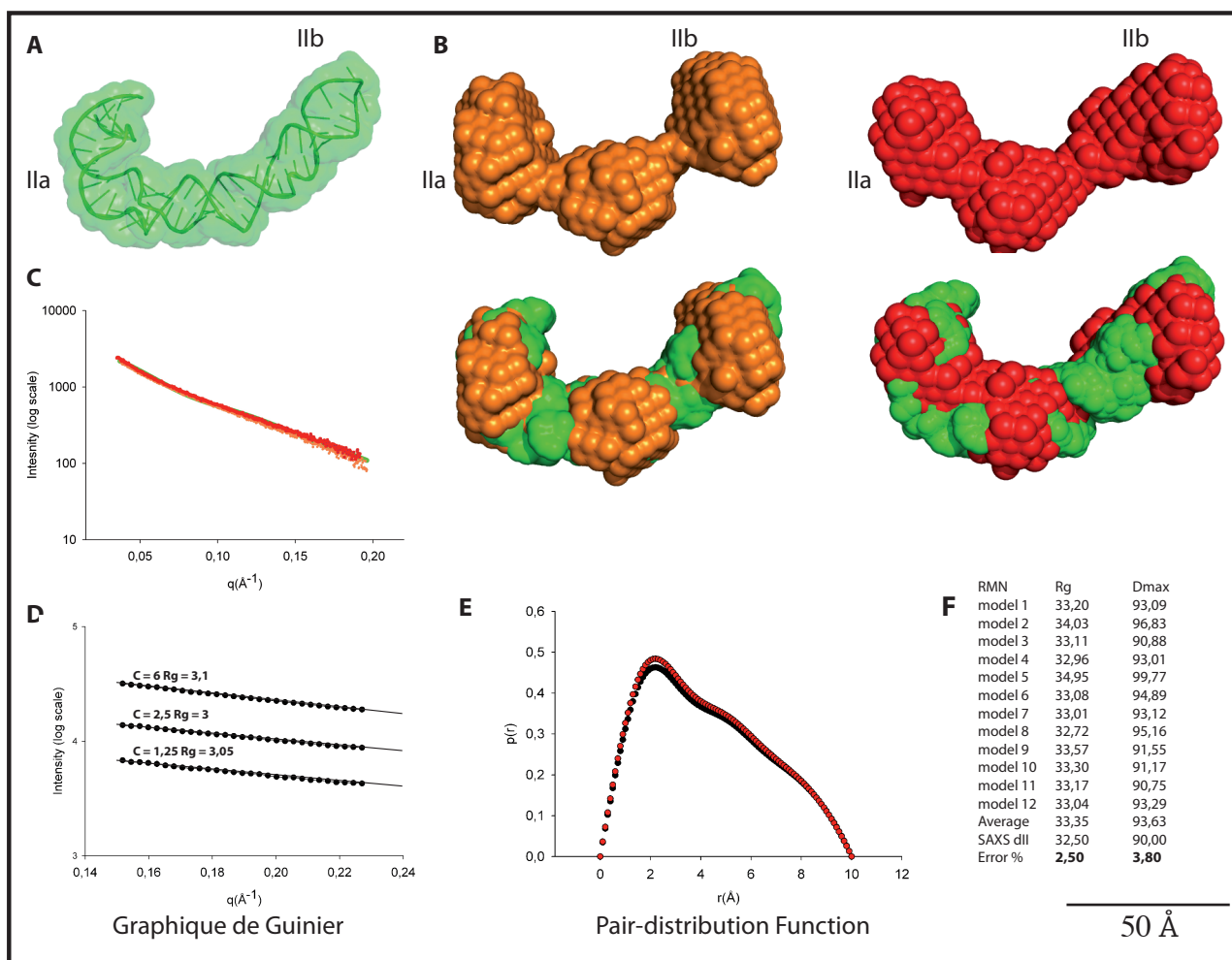


FIG26: Comparaison des données haute résolution et basse résolution sur le domaine II de l'IRES.

A : Modèle haute résolution du domaine II (RMN) (Lukavsky *et al.* 2003).

B : Modélisation du domaine II à partir des données de SAXS. En orange dans le tampon de binding, en rouge dans le tampon utilisé pour l'étude RMN.

C : Superposition des données expérimentales et théoriques du domaine II.

D : Effet de la dilution sur la mesure du Rg par le graphique de Guinier.

E : Superposition de la distribution de fonction des données haute résolution et basse résolution.

F : Données biophysiques de RMN comparées au SAXS.

On note une bonne superposition entre les trois courbes avec un coefficient de corrélation supérieur à 0,98. Lors de l'acquisition de données, j'ai effectué plusieurs mesures de la diffusion du domaine II à plusieurs concentrations (**FIG26D**). Ces mesures m'ont permis de constater l'absence de variation de la mesure des paramètres biophysiques à trois concentrations différentes. J'ai aussi comparé la distribution de fonction entre les modèles « haute » et « basse » résolution (**FIG26E**). Cette représentation graphique nous permet d'apprécier l'état de repliement d'une molécule et de connaître sa forme. Les résultats indiquent que la molécule est de petite taille, de forme globulaire et allongée. Les deux courbes se superposent sur la quasi-totalité du graphique, on note cependant quelques différences, correspondant aux structures « haute » résolution présentes dans le modèle RMN. Dans un dernier temps, j'ai comparé les paramètres biophysiques des douze structures RMN du domaine II disponibles, qui indiquent la présence d'une structure assez stable avec un R_g proche de celui obtenu par SAXS (**FIG26F**). Le domaine II possède deux régions dynamiques situées au niveau de la boucle IIa et dans la partie apicale IIb. Les modèles de SAXS présentent deux constriction, indiquant la présence de régions dynamiques qui pourraient correspondre au deux régions mobiles précédemment identifiées par RMN. Cependant, le modèle de SAXS ne permet pas d'identifier avec précision les régions impliquées. D'après cette analyse, nous pouvons affirmer que la technique de SAXS est bien adaptée à l'étude de molécule d'ARN en solution.

V Construction du modèle de l'IRES entier (II-III-IV)

Les données de SAXS, une fois traitées, permettent d'obtenir des informations sur la forme et l'enveloppe d'une molécule. Plusieurs programmes peuvent ensuite être utilisés pour tenter d'insérer des structures à haute résolution. Dans le cas de l'IRES, qui est une molécule d'ARN et qui possède des domaines flexibles, cette approche est difficile. Tous les programmes disponibles ne permettent pas d'intégrer des données de dynamique moléculaire sur des structures à haute résolution. La forme et la taille du domaine II, qui est une structure rigide, est retrouvée par SAXS comme cela a été décrit dans le paragraphe précédent. Il est cependant possible d'étudier la flexibilité entre deux domaines rigides (technique de « rigid body »). Cependant, ces programmes ne prennent pas en compte la modification de structure à haute résolution. Nous avons utilisé le programme CREDO (Svergun) qui nous a permis de reconstruire le domaine IIIabc en entier grâce aux structures hautes résolutions des domaine IIIb et JuncIIIabc combinés aux données de diffusion du domaine IIIabc. Les résultats sont présentés dans la **figure 27** et nous indiquent la présence des deux structures « haute résolution » dans l'enveloppe du modèle généré, ainsi que l'apparition d'une boucle dans la partie apicale, correspondant aux seize nucléotides manquants. La superposition de la courbe expérimentale et théorique après introduction de la structure correspondant à 66 nucléotides sur 88 (soit 75%) a permis de modéliser la boucle apicale IIIb de l'IRES du VHC. Cette boucle semble former un angle droit et correspond à la forme du domaine IIIb présent dans les modèles des diffé-

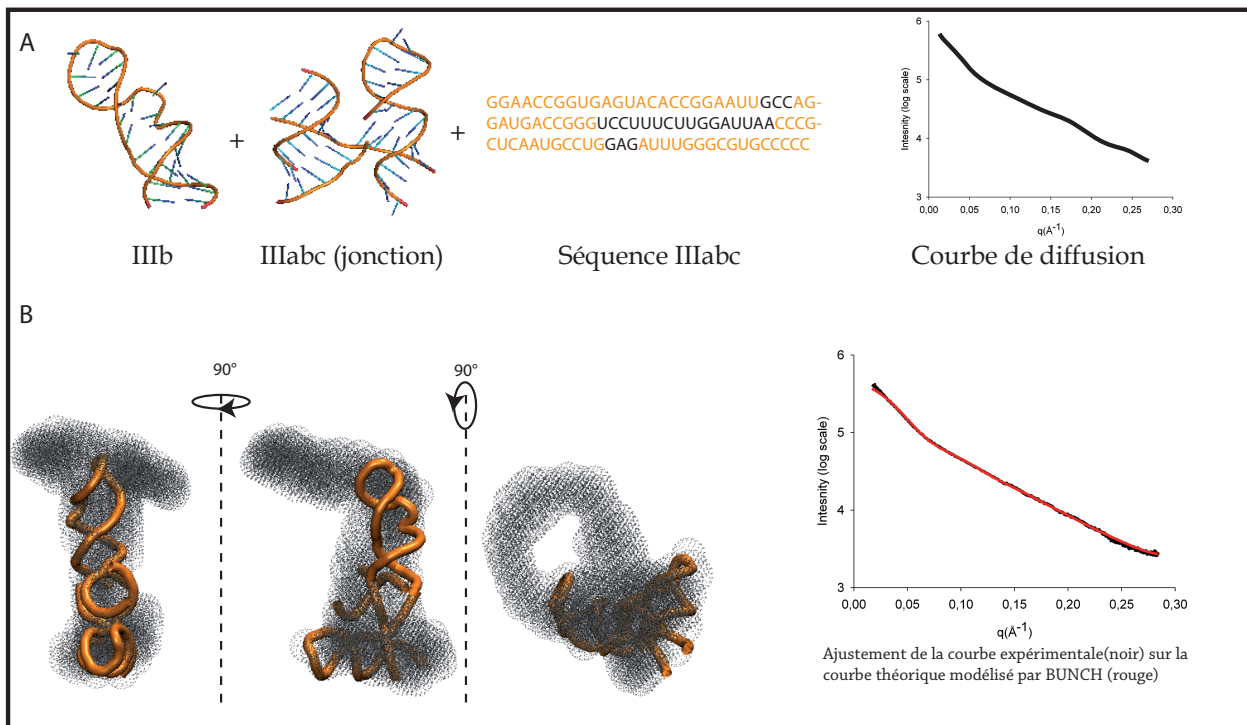


FIG27 : Remplacement moléculaire du domaine IIIabc par le programme CREDO et insertion des séquences haute résolution dans l'enveloppe de SAXS.

A : Préparation de la modélisation : Le programme nécessite les coordonnées correspondant aux structures haute résolution qui doivent correspondre à la séquence du domaine IIIabc (pdb 1KP7) et du domaine juncIIIabc (pdb 1KH6), un fichier présentant la séquence du modèle final et la courbe de diffusion du modèle final.

B : Modélisation : Le programme CREDO va modéliser les nucléotides manquant en utilisant la courbe de diffusion du domaine IIIabc. Les résultats d'ajustement du programme indiquent un coefficient de corrélation de 0,99. Le programme a modélisé une boucle de 16nt pour compléter le modèle actuel.

rents fragments. La technique de SAXS a permis ici d'appréhender la structure d'un domaine très « dynamique » de vingt nucléotides impossible à étudier par les techniques classiques. En utilisant la même approche, j'ai tenté d'introduire toutes les structures « haute résolution » dans chaque enveloppe. Cependant, les structures ne représentant que de petits domaines de chaque fragment étudié, cette approche ne s'est pas révélée efficace. Après avoir modélisé, sans succès, tous les autres fragments par plusieurs techniques, j'ai opté pour une approche de modélisation manuelle. Grâce à la modélisation des domaines individualisés, il est possible de comparer chaque modèle et de localiser chaque sous domaine indépendamment. Par superposition des différents modèles, j'ai reconstruit successivement différents domaines, pour arriver à reconstruire en dernier lieu l'intégralité de l'enveloppe de l'IRES du VHC.

Les modèles IIIabc, IIIabcd et IIIabcdef ont été utilisés pour former le modèle IIIabcdef (**FIG28**). A l'aide du programme «VMD» nous avons réussi à superposer manuellement chaque domaine les uns sur les autres. Tout d'abord, j'ai disposé le modèle IIIabc dans le modèle IIIabcd. On retrouve ainsi la courbure formée par le domaine IIIabc, et on note la présence d'un allongement de la molécule en présence du domaine IIId situé dans la partie inférieure du domaine IIIabcd. J'ai ensuite utilisé ces résultats pour orienter le modèle IIIabcdef ; dans ce cas on retrouve le domaine apical IIIabc avec sa forme en demi-cercle, suivi d'un repliement de la structure pré-

sentant le domaine III d et descendant jusqu'aux domaines III e et III f. A la suite de la génération de ce modèle, j'ai réalisé un modèle simple à deux dimensions me permettant d'apprécier la disposition de chaque sous domaine l'un par rapport à l'autres. Comme cela a été prédit par la structure secondaire, on retrouve les trois régions principales du domaine III, à savoir III abc, III d et III ef.

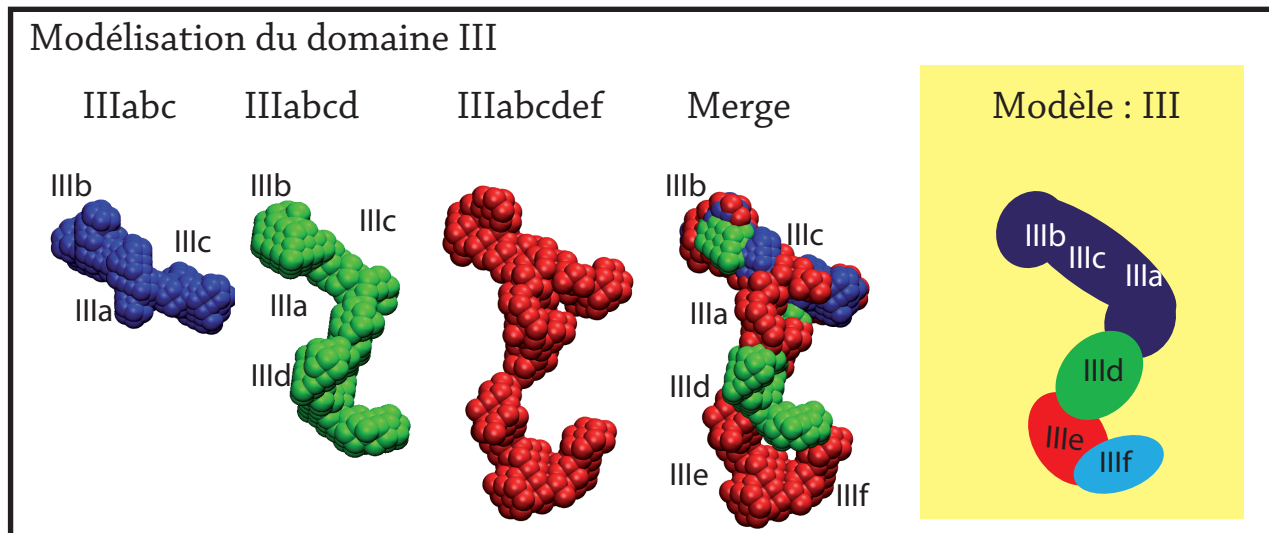


FIG28 : Modélisation manuelle du domaine IIIabcdef.
Les modèles IIIabc, IIIabcd et IIIabcdef sont alignés manuellement à l'aide du programme VMD.

Nous avons poursuivi nos investigations en utilisant la même approche pour reconstruire le modèle II, III (**FIG29**). On peut situer le domaine II dans la partie basale du domaine III. Ces résultats confirment la prédiction de structure, et montrent la présence du domaine II à la perpendiculaire du domaine III. La forme du domaine III ne semble pas être influencée par la présence du domaine II, et l'on retrouve la même orientation des sous domaines. Les données biophysiques confirment que la présence du domaine II n'augmente que très peu le Rg ou le Dmax de la molécule.

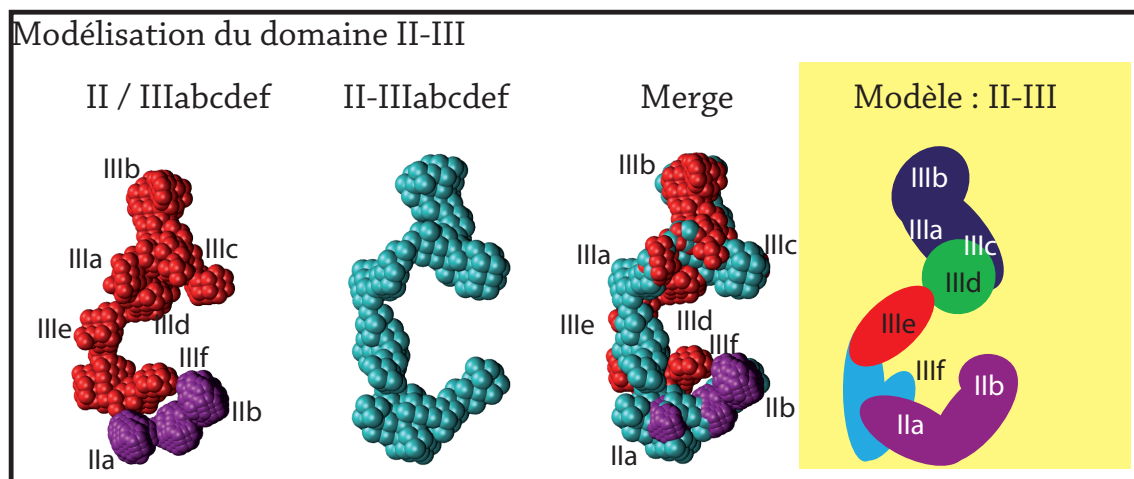


FIG29 : Modélisation manuelle des domaines II-III
Les modèles II, III et II-III sont alignés manuellement à l'aide du programme VMD.

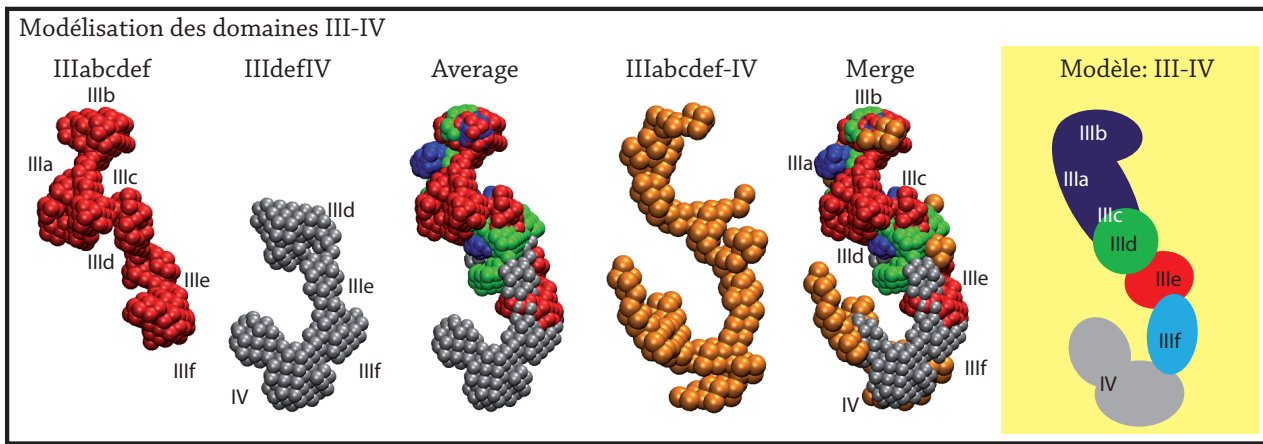


FIG30 : Modélisation manuelle des domaines III-IV

Les modèles III, IIIdefIV et III-IV sont alignés manuellement à l'aide du programme VMD.

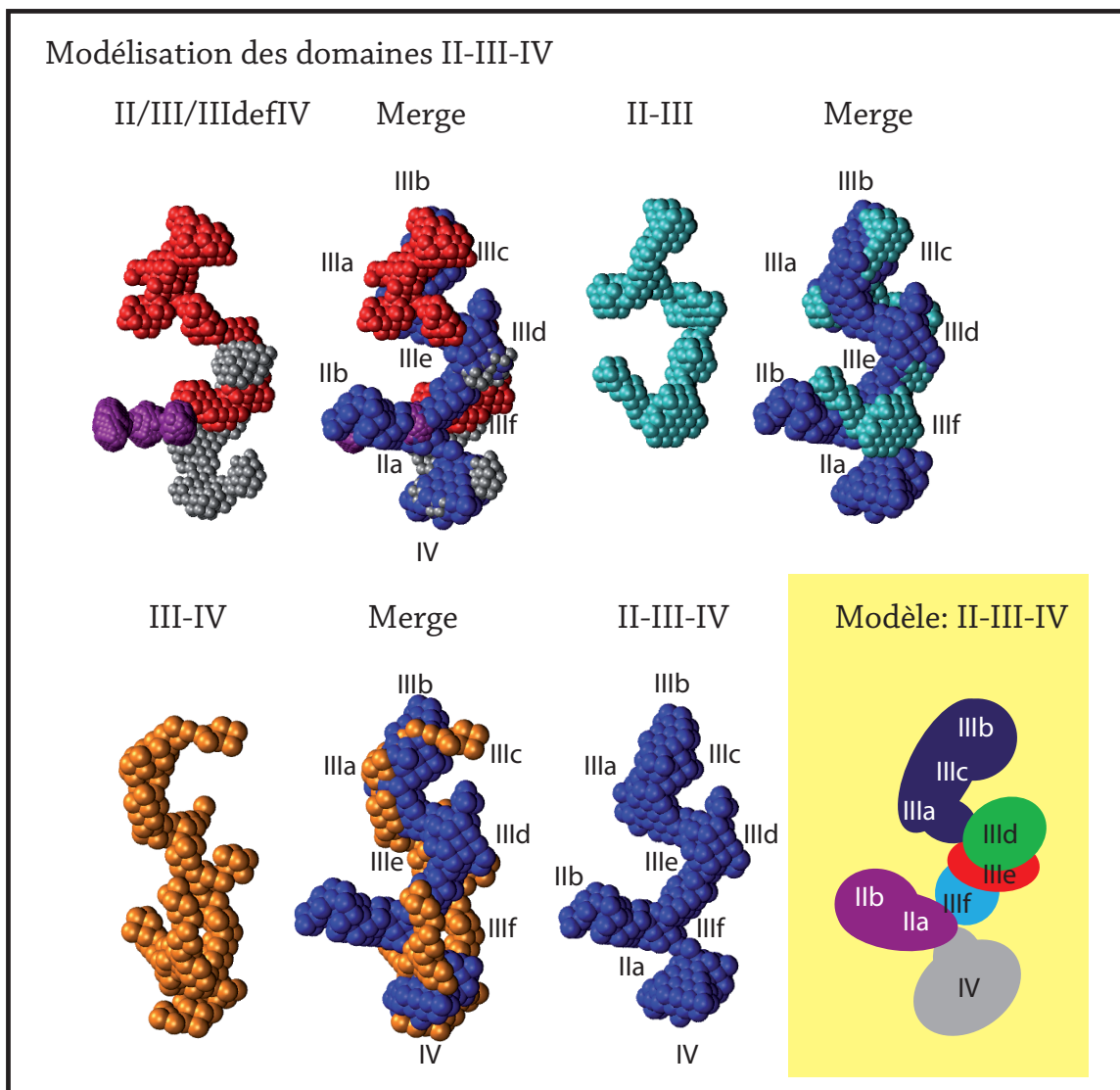


FIG31 : Modélisation manuelle de l'IRES II-III-IV par reconstruction à l'aide du programme VMD. Les modèles IIIabcdef, II-III et III-IV sont superposés au modèle II-III-IV.

En dernier lieu, j'ai étudié le domaine IV constituant le dernier domaine de l'IRES. Pour ce faire j'ai superposé les modèles IIIabcdef, IIIdefIV et III-IV (**FIG30**). Dans ce cas, on peut constater que le domaine IV est situé dans la partie basale de la molécule. On retrouve la forme du domaine III, ce qui nous indique que le domaine IV n'influence pas directement sur le domaine III. Après avoir construit ces modèles et repéré chaque sous domaine les uns par rapport aux autres, j'ai pu envisager de construire le modèle de l'IRES II-III-IV.

VI Présentation de l'enveloppe de l'IRES

En utilisant la même approche présentée dans le paragraphe précédent, nous avons réussi à obtenir l'enveloppe de l'IRES entier composé de ses trois domaines (**FIG31**). Grâce à ce modèle, nous avons pu apprécier chaque domaine de manière individuel au sein de la structure. Le domaine II est localisé à la perpendiculaire du domaine III, et le domaine IV est localisé dans la partie inférieure de l'IRES sous le domaine IIIf. On retrouve la forme de chaque domaine préalablement défini dans les modélisations antérieures. On remarque que la région IIIef permet d'assembler les trois domaines de l'IRES, et oriente le domaine II par rapport aux domaines III et IV. Le domaine II est toujours retrouvé à la perpendiculaire des domaines III-IV qui forme une structure allongée. Ces résultats semblent indiquer que les domaines II, III et IV sont autonomes et n'interagissent pas directement entre eux. Afin de parfaire ce modèle j'ai positionné dans l'enveloppe de l'IRES, les structures tridimensionnelle à hautes résolutions disponibles dans la base de données pbd (domaine II pdb 1P5P, jonction IIIabc pdb 1KH6, domaine IIIb pdb 1KP7, domaine IIIc pdb 1IDV, domaine IIId pdb 1F84, domaine IIIe pdb 1F85) (**FIG32**). On remarque qu'il est possible d'identifier chaque structures dans l'enveloppe générée par SAXS. On retrouve les formes des domaines II, IIIabc ainsi que celle du domaine IIIb. Concernant le

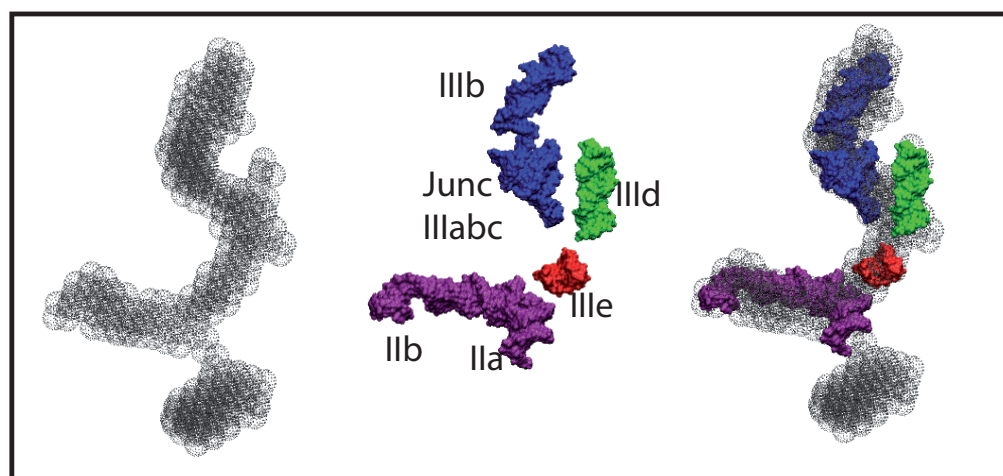


FIG32 : Superposition des structures à haute résolution dans l'enveloppe de l'IRES II-III-IV modélisée par SAXS.

domaine IIIId, celui-ci semble être localisé à proximité immédiate du domaine IIIc situé de ce fait au centre de la molécule.

De précédents travaux de microscopie électronique, obtenus en 2001, avaient permis d'identifier deux zones flexibles localisées entre les domaines II et III ainsi que sur le domaine III (cercles jaunes dans la **figure 33**). Les auteurs de cette étude ont estimé que le domaine II était capable de s'ouvrir de 130° ou de se fermer de 70° par rapport au domaine III. Le domaine II peut donc effectuer un mouvement maximum de 200° par rapport au domaine III. La présence d'une zone flexible située au niveau de la jonction IIIabc a aussi été identifiée (**FIG33**).

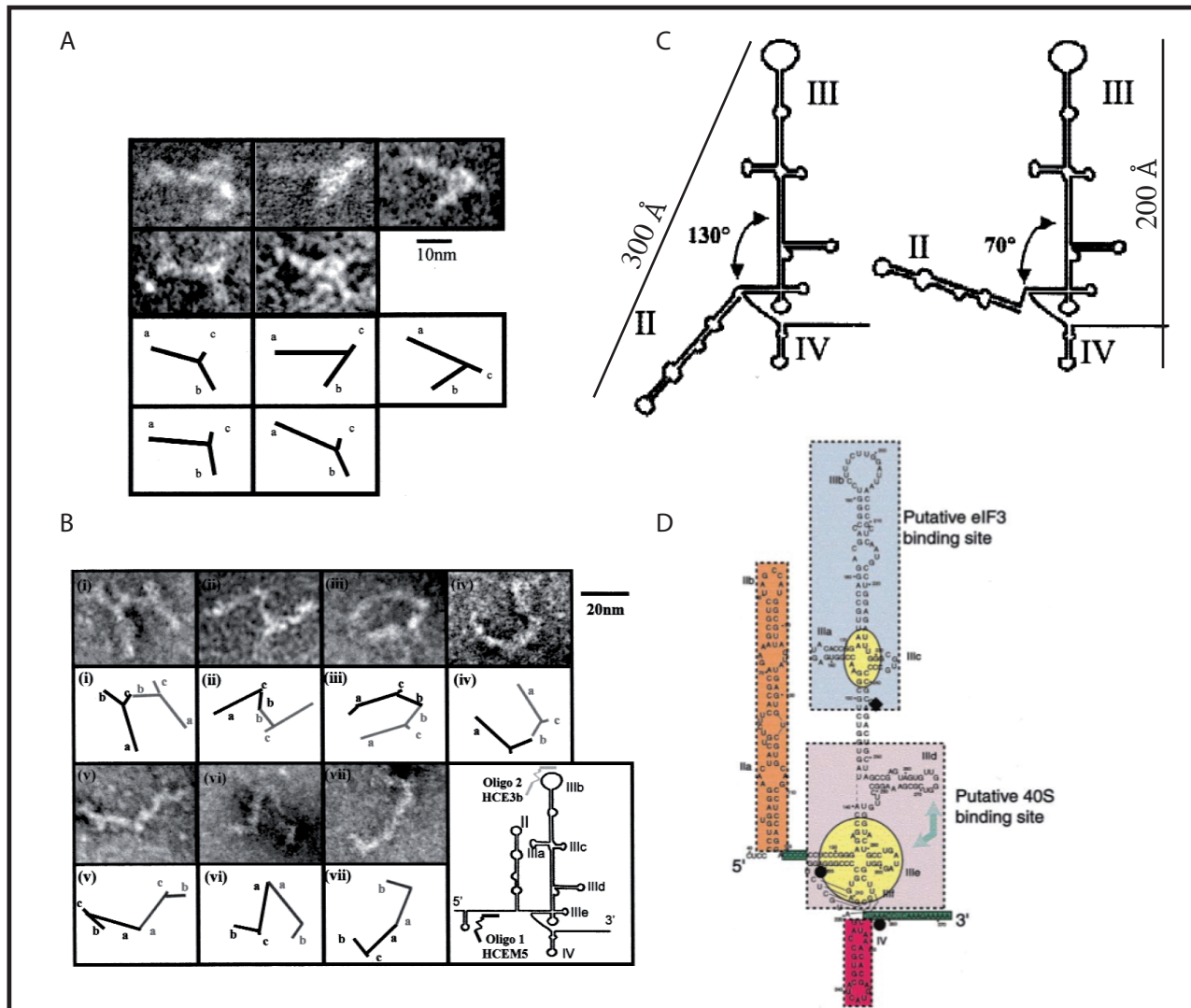


FIG33 : Etude de l'IRES par microscopie électronique et construction d'un modèle 2D.

A : Visualisation par microscopie électronique en coloration négative réalisée avec de l'acétate d'uranyle 4% suivant le protocole de Hoppert and Holzenburg (1998).

B : Visualisation par microscopie électronique en coloration négative après couplage de 2 molécules d'IRES par la technique de «Tandem repeat oligonucleotides».

C : Modèle de l'IRES du VHC basé sur la structure IIIb primaire et les résultats de microscopie électronique. Source : The internal ribosome entry site (IRES) of hepatitis C virus visualized by electron microscopy : Beales et al., 2001.

D : Etude biochimique de l'IRES du VHC et identification du domaine flexible.

En jaune, les domaines flexibles, caractérisés par empreinte chimique en présence d'ions. (Kieft et al., 2001).

Cependant il semble difficile de mesurer avec précision un tel mouvement avec des données de microscopie électronique.

D'autres données de la littérature indiquent que l'IRES, en présence de la sous-unité 40S, possède une forme allongée de 30 nm (Sirideshadilock *et al.* 2005) (**FIG34**). A l'aide de la microscopie électronique les auteurs ont identifié 3 régions distinctes : IIb, III d/e/f et IIIa/c + IIIb. La première région est formée par le domaine II qui est situé dans le sillon de l'ARNm et présente une courbure caractéristique en faucille. La région centrale est formée par le domaine IIIdef. La boucle apicale du domaine III est quant à elle retrouvée libre dans la dernière partie (**FIG34-A**). D'après les données de microscopie électronique étudiant l'IRES en complexe avec 40S, une structure tridimensionnelle à 15Å de résolution a été proposée et présente les domaines IIIb, IIIabc et II (Boehringer *et al.* 2005) (**FIG34B**). D'après ces données, on peut remarquer que le domaine II est retrouvé à l'opposé du domaine III et contribue de ce fait à l'allongement de la molécule. Cette orientation particulière amène le domaine à se positionner dans le sillon de l'ARNm de la sous-unité 40S. On constate également que les domaines III et IV semblent libres et potentiellement accessibles, aux autres facteurs cellulaires comme par exemple eIF3.

La comparaison de l'enveloppe de microscopie obtenue pour le complexe 40/IRES au modèle généré par SAXS nous indique que le domaine II n'est pas dans la même position (**FIG35**). Il semblerait en effet que le domaine II soit orienté différemment à la perpendiculaire du domaine

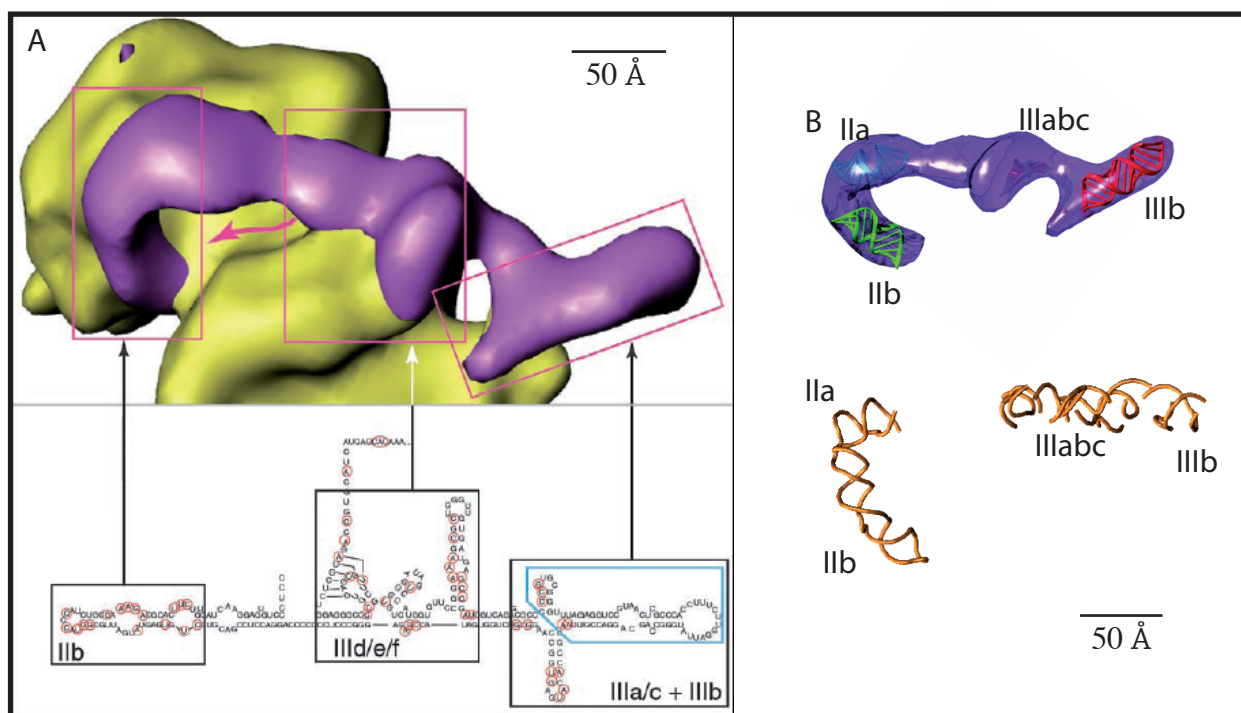


FIG34 : Présentation de l'IRES en complexe avec la sous-unité 40S

A : Visualisation en 3D, après reconstruction d'images obtenues en microscopie électronique du complexe 40S/IRES (Siridechadilok *et al.* 2005).

B : Structure à 15Å de résolution de l'IRES du VHC combinant les coordonnées des structures pdb du domaine II, IIIabc et IIIb.

III, raccourcissant de ce fait la taille de l'IRES de 100Å soit 200Å. Le domaine III semble présenter la même architecture et possède une forme allongée.

En revanche, les domaines II et IV dans le modèle de SAXS sont dans une disposition opposée au modèle de microscopie électronique. Afin de positionner l'IRES en solution dans l'orientation déterminée en complexe avec la sous-unité 40S, il est nécessaire de faire pivoter le domaine II d'un angle de 90° par rapport au domaine III (**FIG35**). Cette modification conduit à l'allongement de 100Å de l'IRES qui se retrouverait alors avec une taille de 300Å (200 + 97 qui provient du domaine II). La combinaison de ces résultats indique qu'il existe une région flexible, localisée entre les 3 domaines de l'IRES, probablement localisée proche du domaine centrale IIIdef, qui conduit à l'ouverture du domaine II et à l'allongement de la molécule lors de son interaction avec la sous-unité 40S. Afin d'étudier la dynamique de cette région j'ai utilisé les données de diffusions de SAXS des différents domaines afin de déterminer l'orientation du domaine II par rapport aux domaines III et IV.

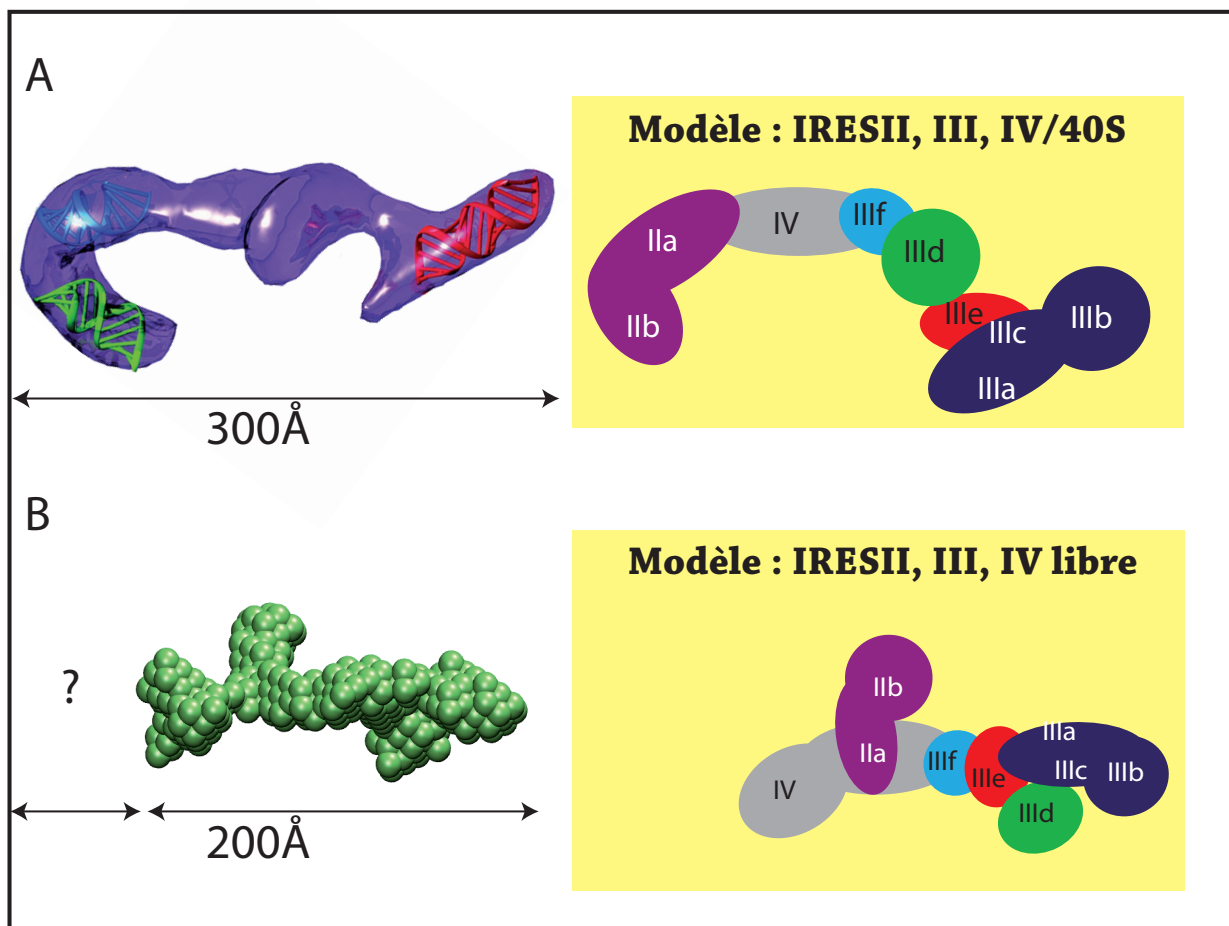


FIG35 : Comparaison des modèles de l'IRES, libre ou en complexe

A : Modélisation de l'IRES en complexe avec la sous-unité 40S (microscopie électronique).

B : Modélisation de l'IRES seul en solution (SAXS).

VII Etude de la dynamique du domaine II

Le domaine II présente une forme particulière, aisément reconnaissable parmi les autres domaines de l'IRES. Les résultats obtenus par SAXS indiquent clairement que ce domaine est situé à la perpendiculaire du domaine III. La forme de l'IRES en présence de la sous-unité 40S est allongée, et situe le domaine II à l'opposé du domaine III à un angle de 180° . Il y a donc un mouvement de d'ouverture de 90° . La jonction entre le domaine II et le domaine III confère donc une certaine flexibilité entre ces deux domaines. Afin de confirmer cette hypothèse et d'appuyer les résultats de microscopie électronique qui indiquent que le domaine II est susceptible de s'orienter et s'ouvrir d'un angle de 180° , nous avons étudié la dynamique du domaine II à l'aide du programme SASREF (Svergun) qui permet d'étudier les mouvements entre deux structures rigides.

J'ai tout d'abord estimé la mobilité du domaine II en présence du domaine III. Pour ce faire, nous avons utilisé l'enveloppe du domaine II-III au sein duquel nous pouvions localiser précisément le domaine II. J'ai identifié la position du domaine II entre les domaines IIIe et IIIf, l'orientant de ce fait dans la position définie par les données de microscopie électronique. Grâce au programme SASREF, j'ai positionné le domaine III dans une orientation me permettant d'étudier la dynamique du domaine II. J'ai ensuite indiqué au programme le degré de liberté du domaine II en lui permettant de bouger d'un angle de 200° et de réaliser un mouvement maximal de 200\AA à partir de sa position initiale. J'ai réalisé 4500 modélisations SASREF afin d'établir une bonne répartition statistique des différents mouvements possibles du domaine II par rapport au domaine III. A la suite de cette première étape, qui a permis de générer des modèles contenant le domaine II dans un grand nombre de positions, j'ai « compilé » ces résultats à l'aide du programme EOM (Ensemble Optimisation Modeling). Ce programme nous a permis de comparer tous les modèles entre eux, à l'aide de la courbe de diffusion du domaine II-III. L'analyse des résultats a permis d'obtenir quatre orientations différentes du domaine II avec un mouvement uniquement horizontal d'un angle maximum de 45° (**FIG36**). On peut donc constater que le domaine II ne peut pas s'orienter à 180° par rapport au domaine III dans le domaine II-III mais seulement de 90° . L'absence du domaine IV peut expliquer ce résultat et indiquerait que l'intégralité de l'IRES est requis pour effectuer cette rotation. Afin de poursuivre l'étude de la dynamique du domaine II, nous avons utilisé la même approche pour mesurer la mobilité du domaine II en présence du domaine IV. Pour ce faire, j'ai utilisé le modèle III-IV sur lequel j'ai fixé la position du domaine II. En se basant sur les données de diffusion de l'IRES complet (II-III-IV), et après utilisation du programme SASREF et compilation des résultats par EOM, j'ai pu visualiser les mouvements du domaine II (**FIG36**). Les résultats nous indiquent que le domaine II peut cette fois s'orienter à plus de 180° par rapport au domaine III à l'aide d'un mouvement vertical. Le domaine II peut donc s'ouvrir par rapport au domaine III en présence de l'intégralité de l'IRES (**FIG37**). D'après nos résultats on constate que les domaines

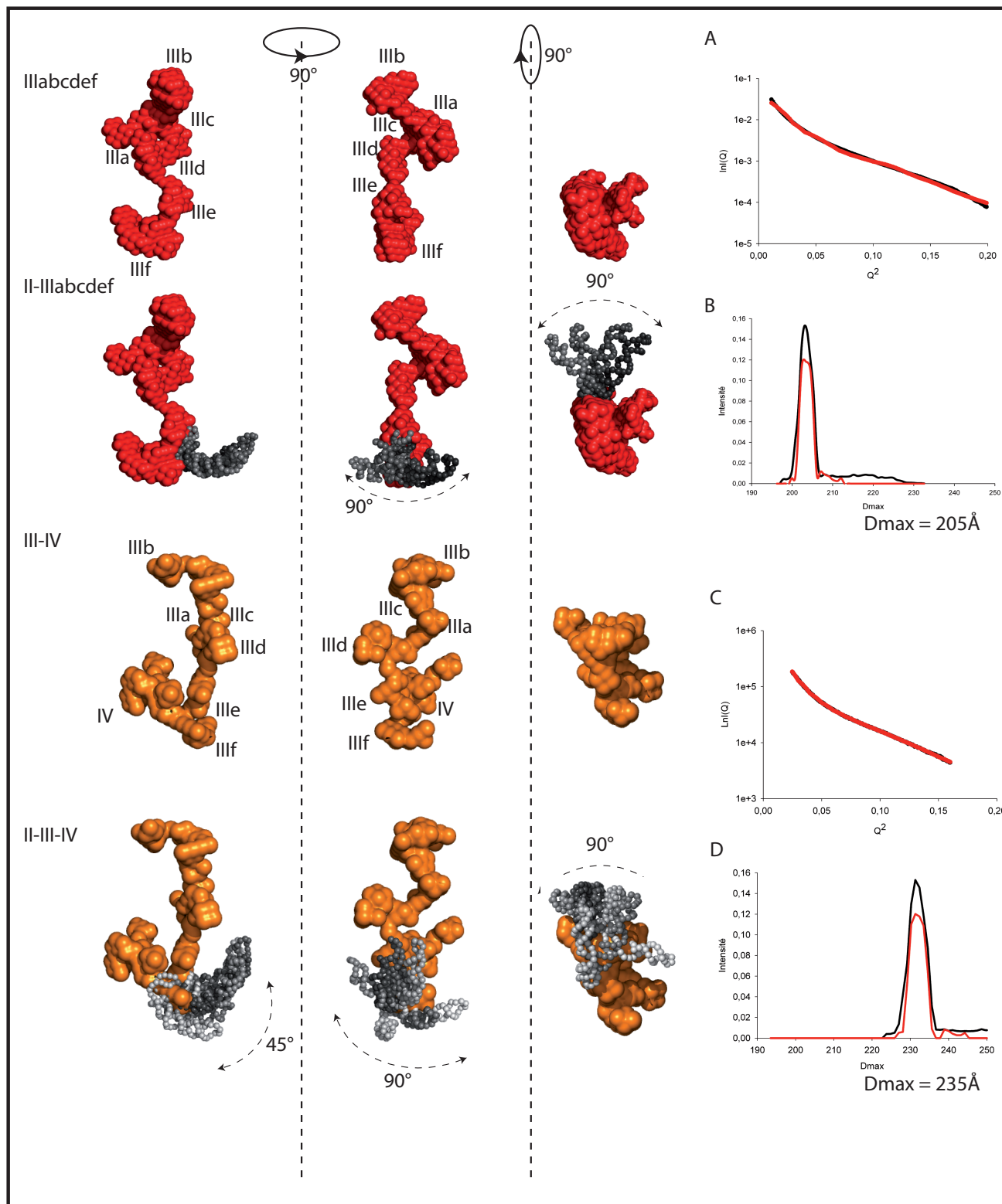


FIG36 : Modélisation de l'étude menée sur la dynamique du domaine II

Superposition des données expérimentales et théoriques après traitement par le programme EOM.

A : Ajustement de la courbe théorique (rouge) sur la courbe expérimentale (noire).

B : Sélection des modèles en fonction du Dmax par EOM.

C : Sélection des modèles en fonction du Rg par EOM.

En noir, distribution statistique des modèles générés par SASREF.

En rouge, sélection des modèles en fonction de la courbe par EOM.

D : Modélisation de la dynamique du domaine II sur le domaine III de l'IRES du VHC.

L'angle maximal du mouvement horizontal du domaine II est de 45°.

n'interagissent pas directement entre eux, mais la présence de chaque domaine apporte la rigidification et/ou la mobilité à certaines régions de la molécule. Cette dynamique particulière confère à l'IRES la mobilité nécessaire à l'interaction avec les différents acteurs protéiques de l'initiation. Ces données confirment de précédents résultats obtenus par Kieft en 2001 (Kieft *et al.* 2001) qui indiquaient que le domaine II était capable de s'orienter de 130° par rapport au domaine III et de ce fait participant à l'allongement de la molécule de 200 à 300\AA .

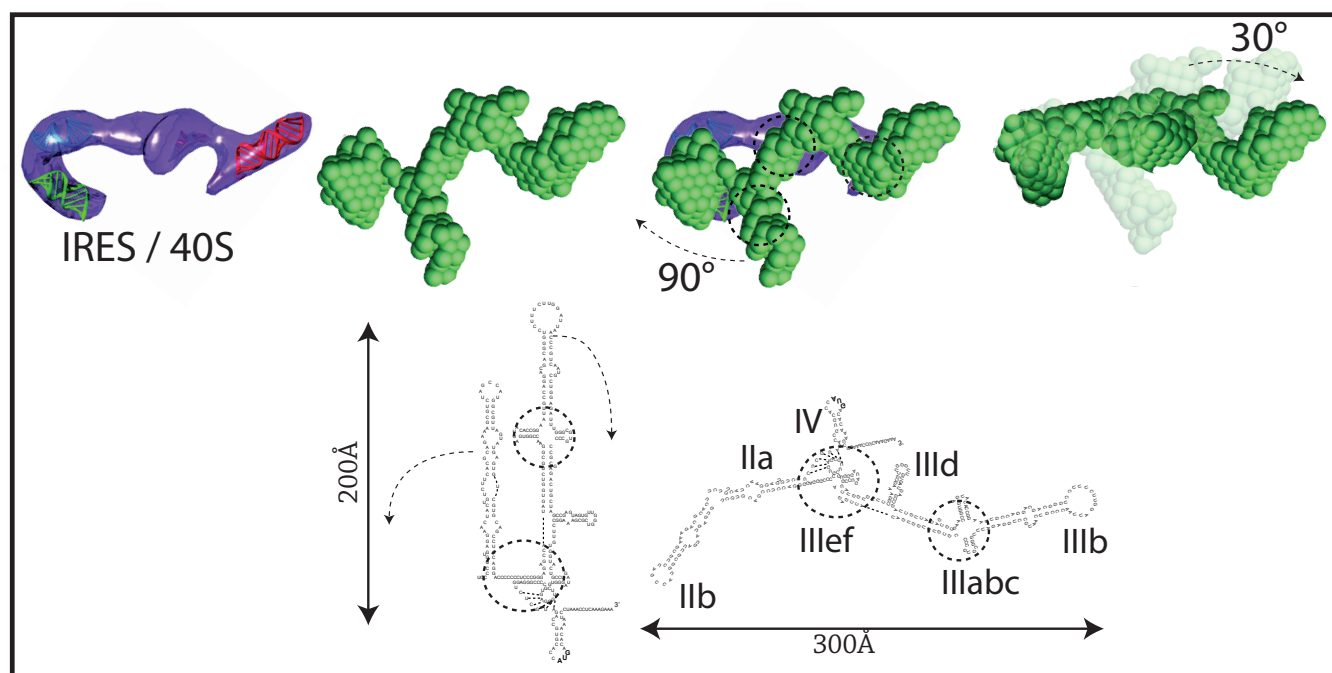


FIG37 : Comparaison de la forme globale de l'IRES, lié à 40S ou absence de facteurs.

VIII Conclusion

D'après les données de la littérature, l'affinité de la sous-unité 40S (2nM) pour l'IRES est supérieure à celle mesurée pour eIF3 (35nM) (Otto *et al.* 2004). On peut penser que la sous-unité 40S est très certainement le premier acteur à se lier à l'IRES du VHC. De ce fait, la fixation de 40S sur l'ARN entraînerait un changement de conformation de l'ARN qui permet de recruter d'autres facteurs. Ce changement de conformation affecte principalement le domaine II qui doit tout d'abord s'orienter à 180° par rapport au domaine III pour se retrouver proche du sillon de l'ARNm de 40S (Fraser *et al.* 2009). Cette modification conduit à l'ouverture de la structure et à la présentation du domaine III qui pourrait alors participer au recrutement du facteur eIF3.

La technique de SAXS s'est révélée être une bonne approche structurale pour appréhender la dynamique de l'IRES ce qui nous a permis d'étudier pour la première fois en détails l'IRES du VHC en solution.

Caractérisation structurale de la
protéine eIF3b et de son motif
MRR

Introduction

Le complexe eIF3 est composé de douze sous-unités dont cinq (a b c d f) ont été décrites comme impliquées dans l'interaction avec l'IRES du VHC (Sizova *et al.* 1998). Parmi ces sous-unités, la protéine eIF3b est présentée comme participant à l'échafaudage du cœur du complexe eIF3 (Zhou *et al.* 2008). Il existe peu d'informations biochimiques sur cette protéine. Elle est retrouvée dans le cœur du complexe eIF3 et interagit avec les sous-unités eIF3g, i, c, a et e (**FIG38**). Cette protéine de 100kDa comporte sept sites de phosphorylation (Damoc *et al.* 2007) tous localisés dans la partie N-terminale de la protéine et contient un motif identifié comme Motif de Reconnaissance à l'ARN (MRR) (aa 185-268), un domaine impliqué dans la liaison à eIF3e (aa 124-413) et un domaine, identifié récemment, impliqué dans la liaison à eIF3j (aa 170-274) (**FIG39**). eIF3b a précédemment été produite en cellules d'insectes (Masutani *et al.* 2007) ce qui a permis la reconstitution du facteur eIF3 *in vitro*. Les informations structurales concernant la protéine eIF3b humaine ne sont que fragmentaires, seule 10% de la structure est connue. Pourtant cette protéine interagit avec cinq autres protéines ainsi qu'avec de l'ARN viral, elle doit donc posséder une structure particulière. L'étude de sa structure permettrait de mieux comprendre son agencement au sein du facteur eIF3 et d'étudier plus en détail son interaction avec l'IRES. Toutes ces informations indiquent qu'il s'agit d'une protéine complexe et la présence de modifications post traductionnelles n'est pas en la faveur d'une production aisée chez la bactérie.

Afin d'aborder la production de cette protéine, pour en étudier sa structure et ses paramètres biophysiques, j'ai introduit la séquence codante de la protéine humaine dans trois vecteurs d'expression procaryote (pET-TEV, pGEX-TEV et pBADM40). En utilisant une souche de bactérie comportant des ARNt supplémentaires des acides aminés R-I-L, j'ai pu exprimer de petites quantités de protéine. Les conditions de culture et d'expression restant assez médiocres nous avons décidé de rechercher d'éventuels domaines solubles et stables par utilisation de la technique de protéolyse ménagée. Cependant, après plusieurs tentatives, je n'ai pas été en mesure d'identifier de domaines stables.

Ayant toujours l'intention d'identifier des domaines de cette protéine pour en faciliter l'étude structurale, j'ai procédé à une dissection moléculaire en me basant sur la prédiction de structure secondaire. Pour ce faire, nous avons utilisé un programme, «MeDor» (Metaserveur Disorder) qui rassemble plusieurs serveurs de prédiction de structure secondaire (**FIG40A**). Les résultats obtenus, nous indiquent que la protéine semble partiellement structurée et composée de 30% de feuillet β . Cette analyse m'a permis d'identifier plusieurs régions pouvant correspondre à d'éventuels domaines solubles. Afin de découper la protéine en fonction des ces prédictions, nous avons utilisé la stratégie suivante : la partie N-terminale de la protéine ne semblant pas structurée selon les prédictions, la suppression des 132 premiers aa de la protéine nous permet d'obtenir une première construction eIF3V1.

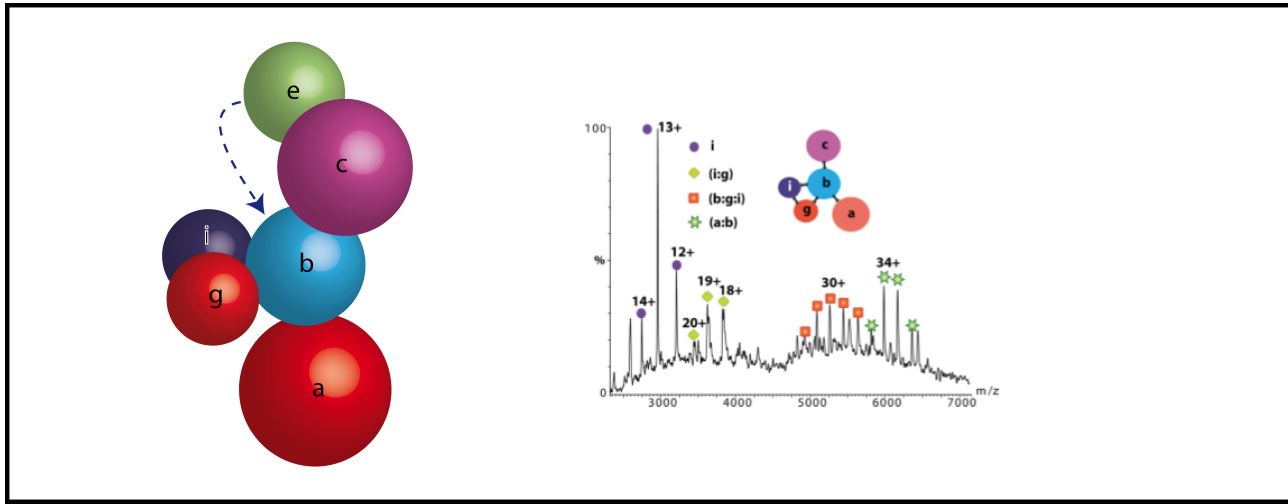


FIG38 : Présentation des interactions de la protéine eIF3b dans le facteur eIF3.
(Inspiré de Zhou *et al.* 2008).

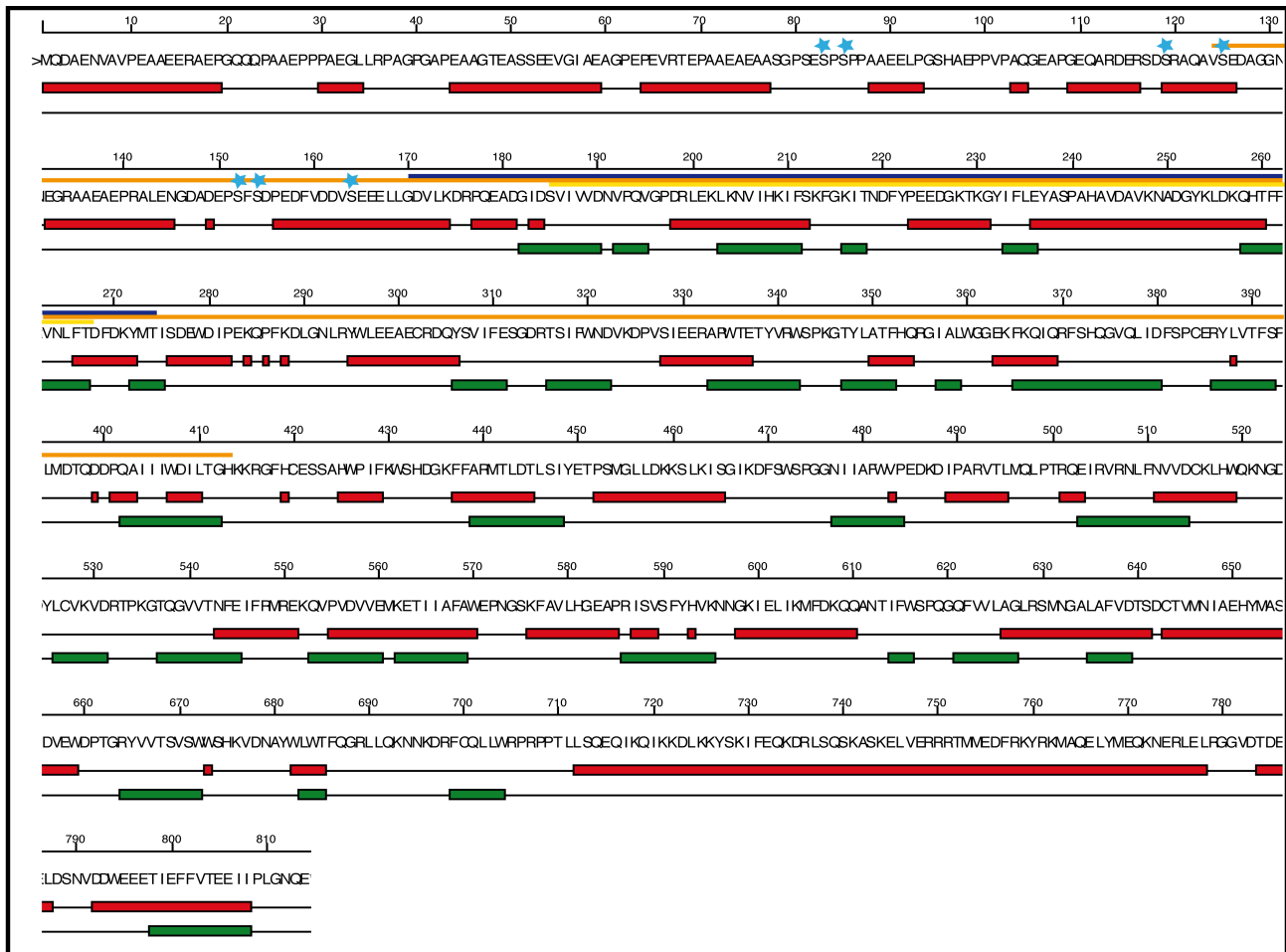


FIG39 : Présentation de la protéine eIF3b

Prédiction de structure secondaire de la protéine eIF3b humaine. (Utilisation du programme LASERGENE de DNASTAR).

Rouge : Hélice α . Vert : Feuillet β . Jaune : motif MRR. Orange : Interaction avec eIF3e. Bleu : Interaction avec eIF3j, étoile bleue site de phosphorylation.

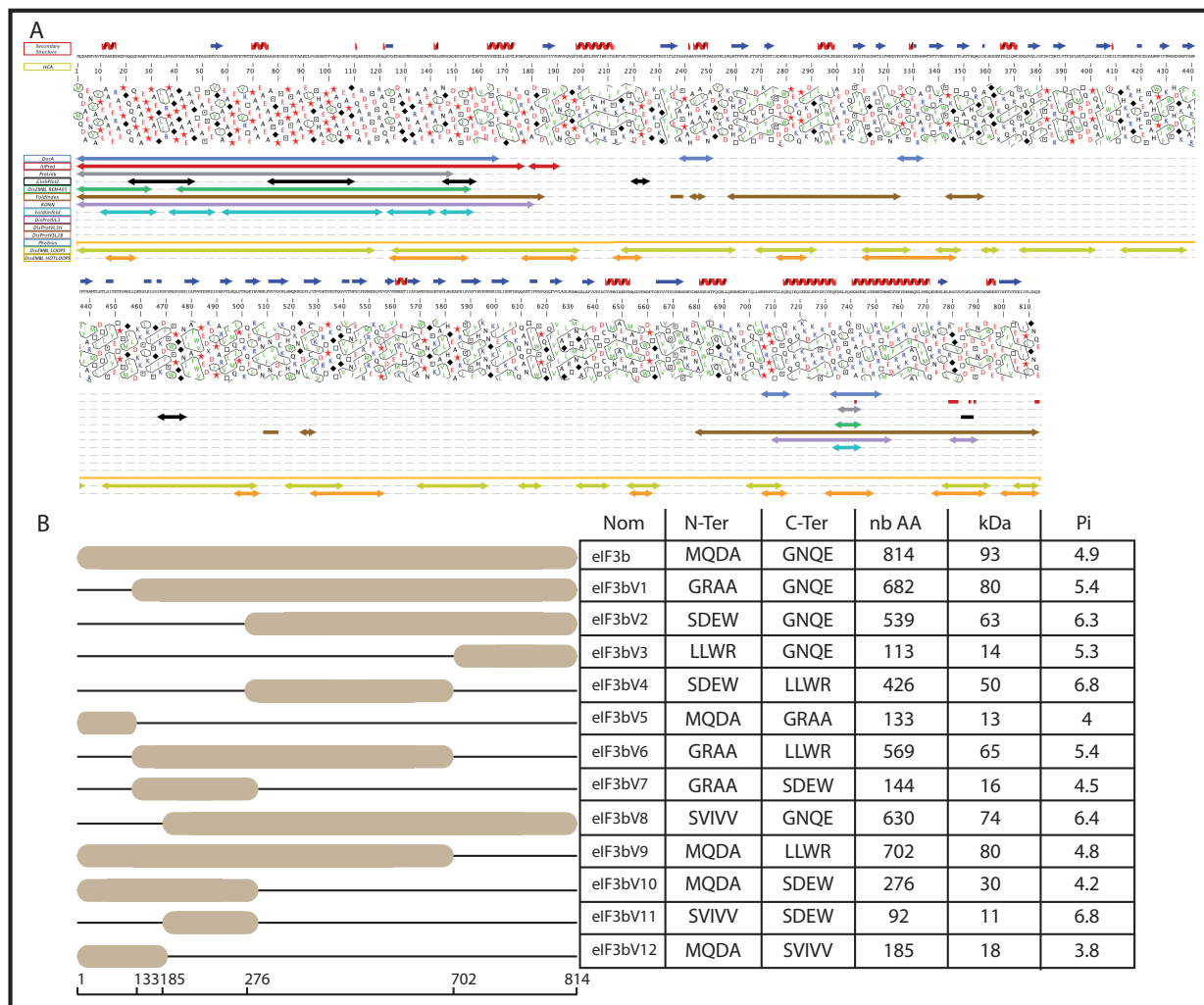


FIG40 : Etude bioinformatique de la protéine eIF3b et identification de sous domaines.

A : Prédiction de désordre et de structure secondaire de la protéine eIF3b humain obtenu par utilisation du programme MeDor. L'analyse des résultats indique la présence de zone non ordonnée en Cter et en Nter de la protéine. On peut remarquer que la partie centrale de la protéine est prédite comme structurée.

B : Dissection moléculaire de la protéine et présentation des données biochimiques de chaque sous domaines.

Afin d'étudier la totalité de la protéine comme de ces domaines, j'ai procédé à la construction complémentaire eIF3v5 qui correspond uniquement aux 132 premiers aa de la protéine. En utilisant la même stratégie, douze constructions ont pu être générées (**FIG40B**). Chaque domaine a été cloné par la technique LIC (Ligation Independent Cloning) dans trois vecteurs d'expression procaryote (pETTEV, pGEXTEV et pETM40). Afin de tester l'expression et la solubilité de ces 36 constructions (12 domaines dans 3 vecteurs), j'ai utilisé un milieu de culture auto-inducible (AI). Cependant, ici aussi, il n'a pas été possible d'identifier de domaines stables. Cette approche après dissection moléculaire de la protéine ne m'a pas permis d'identifier de fragments solubles, et confirme les difficultés rencontrées lors de la recherche de potentiels domaines stables par la technique de protéolyse ménagée. Ayant toujours la ferme intention d'étudier cette protéine, j'ai focalisé mon travail sur le clone exprimant la protéine entière.

Conditions	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Tris 20mM 7.5																				
NaCl 250mM																				
KCl 250mM																				
Sulfate d'Ammo 500mM																				
DTT 1mM																				
BME 5mM																				
Triton 0,1%																				
EDTA 1mM																				
glycérol 5%																				
MgCl2 20mM																				
Conditions	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
PBS																				
NaCl 250mM																				
KCl 250mM																				
Sulfate d'Ammo 500mM																				
DTT 1mM																				
BME 5mM																				
Triton 0,1%																				
EDTA 1mM																				
glycérol 5%																				
MgCl2 20mM																				

FIG41 : Matrice de stabilisation.

Présentation de la composition de la matrice de stabilisation de la protéine eIF3b rassemblant les 40 conditions testées qui sont constituées de 2 tampons, de 3 sels, de 2 agents réducteurs, ainsi que du triton de l'EDTA, du glycérol, du magnésium ou encore du triton X-100.

I Obtention de la protéine eIF3b humaine sous forme recombinante

Après plusieurs phases d'optimisation des conditions de culture et d'expression, j'ai pu obtenir quelque mg de protéine eIF3b humaine recombinante avec la construction pET-TEV (à partir de 12L de culture en milieu LB). Cette protéine a été purifiée sous forme soluble par deux étapes de chromatographies : chromatographie d'affinité au nickel (HisSelect SIGMA) suivie d'une chromatographie de filtration sur gel (S200 GE Healthcare). Cependant, la protéine ainsi purifiée n'est pas stable dans le temps, même en présence de concentration élevée de glycérol (+50%) et/ou d'anti-protéase. Dans le but d'augmenter la stabilité de cette protéine, nous avons utilisé une matrice de « stabilisation » visant à déterminer des conditions de tampon plus adaptées à cette protéine (**FIG41**). Cette matrice a été réalisée en fonction des différents paramètres biochimiques connus de la protéine (pHi, nombre de cystéine...) et m'a permis de définir quarante conditions différentes. Afin de cribler efficacement ces conditions, de petites quantités de culture ont été lysées individuellement dans chacun des 40 tampons. Après centrifugation, la phase soluble est analysée sur gel SDS PAGE. L'analyse des résultats permet de noter un effet stabilisant du tampon phosphate en présence de 250mM de NaCl et de 5mM de β ME (conditions 23 et 25). Ces conditions augmentent la stabilité de la protéine mais elle est toujours sujet à d'importantes dégradations. Afin de rechercher les conditions encore plus adaptées, nous avons, en parallèle de cette technique manuelle, réalisé une étude de « Thermal Shift Assay » (TSA) automatisée permettant de tester un grand nombre de conditions 96 (Laboratoire de D.Hart CIBB-Grenoble). Cette technique utilise les propriétés d'un fluorophore pour estimer l'état de repliement d'une protéine. La plupart des protéines sont sensibles à l'augmentation de la température elles se dénaturent progressivement et exposent des zones hydrophobes. La

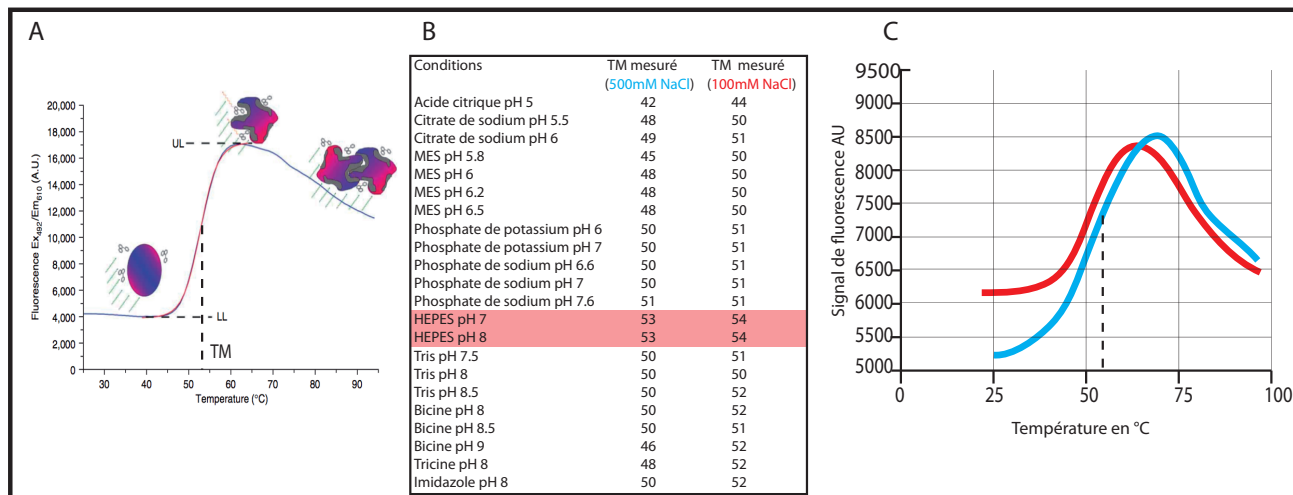


FIG42 : Courbe théorique et étude de stabilité de la protéine eIF3b par thermal shift assay (TSA).

A : Principe du TSA, la variation du signal de fluorescence en fonction de la température permet d’apprécier la stabilité de la protéine. Source : Frank H Niesen, Helena Berglund & Masoud Vedadi.

B : Tableau récapitulatif des TM obtenus sur eIF3b dépendant des conditions utilisées.

C : Exemple de courbes de température obtenu après stabilisation de la protéine en présence de tampon HEPES.

Courbe rouge : eIF3b dans 100mM NaCl 20mM HEPES pH 7

Courbe bleue : eIF3b dans 500mM NaCl 20mM HEPES pH 7

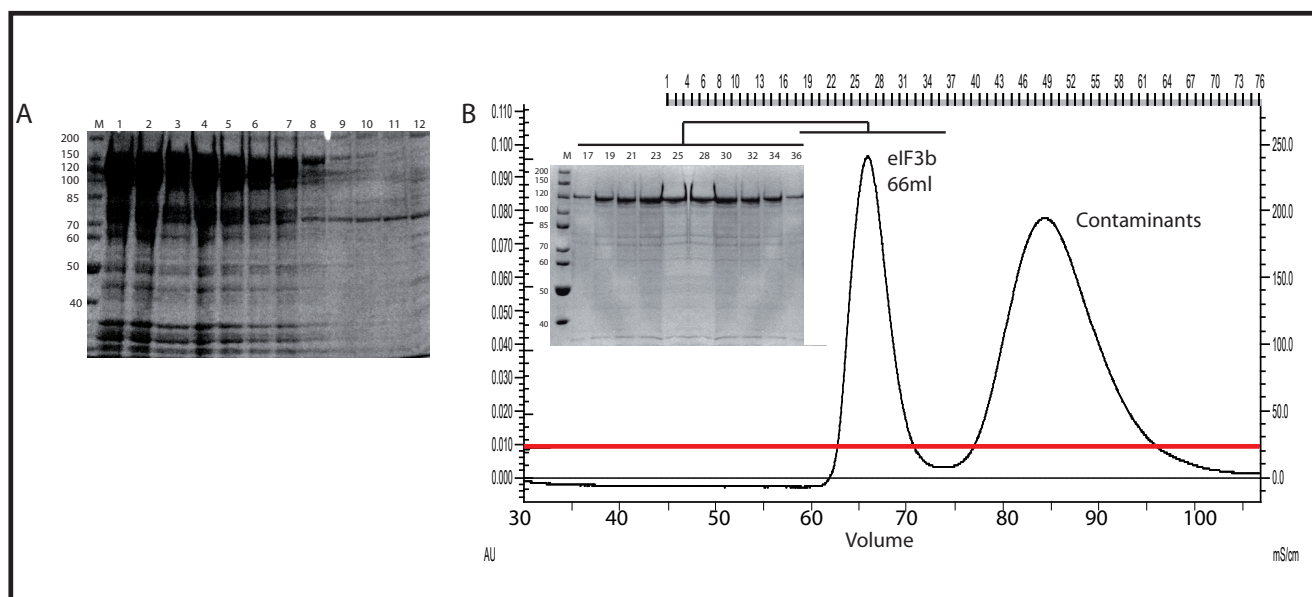


FIG43 : Purification de la protéine eIF3b sous forme recombinante.

A : Gel SDS PAGE 8% représentant les fractions obtenues après la purification par chromatographie d’affinité au nickel (His Select SIGMA) à partir de 2L de culture en AI. Puits de 1 à 12 : fractions d’élution.

B : Gel SDS PAGE 8% et chromatogramme représentant les fractions obtenues après la purification par chromatographie de filtration sur gel S200. Tampon : 20mM HEPES pH 7.5, 150mM NaCl, 5mM BME, 5% glycérol. M : Marqueur de poids moléculaire. 17 à 36 : fractions d’élution analysées. Le volume d’élution du pic est de 67ml ce qui correspond à une taille de 100kDa pour une taille réelle de 96kDa pour la forme contenant un TAG Histidine.

sonde fluorescente (sensible aux zones hydrophobes) va se fixer sur les protéines en cours de dénaturation. Plus la protéine est dépliée, plus elle expose de surfaces hydrophobes, et plus la sonde pourra se fixer. Les résultats sont représentés par un graphique qui traduit l'augmentation du signal de fluorescence en fonction de la température. L'excitation est effectuée à 492 nm et la mesure de l'absorbance est réalisée à 610 nm (**FIG42**). La température de fusion (TM) est calculée en mesurant la température correspondant à 50% du signal maximum de fluorescence. L'analyse des résultats a permis de visualiser l'effet stabilisant du tampon HEPES à pH 7 ou 8 ainsi que la présence de NaCl qui se caractérise par un TM de 53°C (courbe bleu 500 mM NaCl, courbe rouge 100 mM NaCl). Dans la suite des études menées sur eIF3b, j'ai décidé d'utiliser le tampon suivant : 20mM HEPES pH 7; 150 mM NaCl; 5% de glycérol et 10 mM β ME. Grâce à ces nouveaux paramètres, j'ai pu extraire plus de protéine sous forme soluble et obtenir 15 mg de protéine relativement stable à partir de 2L de culture (**FIG43**). Par la suite, nous avons caractérisé cette protéine avant de débiter une campagne de cristallogenèse. En utilisant la technique de dichroïsme circulaire (CD) et de fluorescence intrinsèque des tryptophanes, nous nous sommes assurés de l'état de repliement de la protéine ainsi que de la nature de sa structure secondaire.

II Caractérisation de la protéine eIF3b et approche structurale

Le CD est une technique rapide utilisée en routine au laboratoire, permettant la caractérisation structurale de molécule biologique. Son principe repose sur l'interaction entre une lumière circulairement polarisée et des molécules asymétriques (retrouvées généralement dans des échantillons biologiques). La forme du spectre CD permet de déterminer le contenu en structures secondaires de la molécule étudiée. Pour ce faire, on mesure l'absorbance d'un échantillon avec de la lumière polarisée successivement à droite et à gauche et on observe une différence de l'absorbance $\Delta A = A_{\text{gauche}} - A_{\text{droite}}$. Avec un trajet optique « l » et une concentration « C », le pouvoir dichroïque est considéré comme étant la somme des mesures effectuées à gauche et à droite. Dans le cas des protéines, les éléments de structure secondaire possèdent un pouvoir dichroïque particulier : une hélice α (3,6 résidus/tour) possède un pouvoir dichroïque négatif à 222 nm et 208 nm. Les feuillet β possèdent un pouvoir dichroïque négatif plus faible, à 211 nm. Les entités non structurées possèdent une absorption caractéristique négative. **La figure 44A** présente les résultats obtenus et montre la courbe expérimentale (noire) mesurée pour la protéine eIF3b (5 μ M) à l'aide d'un spectropolarimètre JASCO J-810 et d'une cuve en quartz de 1 mm de trajet optique. Le spectre obtenu est caractéristique d'une protéine composée de feuillet β avec un maximum mesuré à 211 nm de -15 (**FIG44B**). La courbe obtenue est très proche de la courbe théorique (en bleu) qui représente une protéine constituée d'une majorité de structures en feuillet β . Les courbes rouge et verte correspondent respectivement à

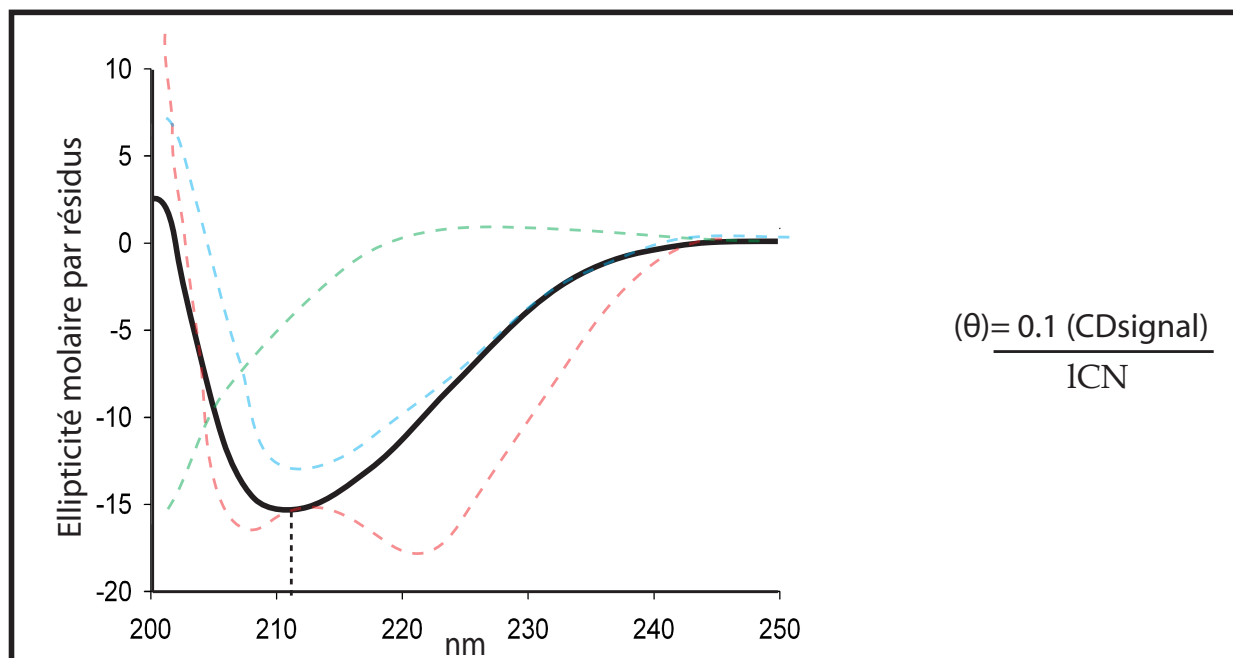


FIG44 : Analyse de la structure secondaire de la protéine eIF3b par dichroïsme circulaire (CD).

Principe du CD et présentation du spectre théorique. En rouge, spectre caractéristique de l'hélice α : pouvoir dichroïque négatif à 222 et à 208nm, pouvoir dichroïque positif à 190nm. En bleu, spectre caractéristique des feuillets β : pouvoir dichroïque négatif à 211nm. En vert, spectre caractéristique de l'absence de structure : pouvoir dichroïque négatif à 190nm.

Equation permettant le calcul de l'ellipticité à partir des données de CD. θ : Ellipticité. CD signal : donnée de CD brut. l : taille de la cuve en cm. C : concentration de la protéine en molaire. N : nombre d'acide aminé de la protéine

des protéines riche en hélice α ou non structurées et ne correspondent pas à la courbe mesurée pour la protéine eIF3b. Avec les données de CD, il est possible d'estimer le % de structure en hélice α présent dans la protéine eIF3b en utilisant les données de l'ellipticité molaire mesurée à 211nm (J.M.Scholtz et R.L Baldwin. 1992). Les auteurs ont déterminé qu'une protéine composée uniquement d'hélice α possédait une l'ellipticité molaire spécifique de -35000. Les données mesurées pour la protéine eIF3b sont $(-0,1*15)/(0,1*0,000005*814) = -3250$, ce qui correspond à environ 9% d'hélice α . Ces résultats sont compatibles avec la prédiction de structure secondaire et indiquent que la protéine est formée majoritairement de feuillets β . Afin de confirmer ces résultats et de m'assurer du bon repliement de la protéine, j'ai utilisé la spectrophotométrie de fluorescence des tryptophanes. C'est une technique rapide qui permet de visualiser l'état de repliement de protéine contenant des tryptophanes. La fluorescence intrinsèque des résidus tryptophane (Trp) de la protéine est sensible à la polarité de l'environnement, et permet de suivre les changements de conformations des protéines. Après excitation à 295 nm, les Trp émettent à 350 nm s'ils sont dans un environnement polaire (c'est à dire exposés au solvant), ce qui correspond généralement à un état déplié. Au contraire, les Trp émettent à 330 nm s'ils sont dans un environnement apolaire et donc dans des domaines structurés. Le spectre de fluorescence (noir) réalisé sur la protéine eIF3b (20 μ M) indique un maximum à 335 nm, ce qui correspond à un état replié en comparaison du spectre théorique (rouge) qui posséderait un

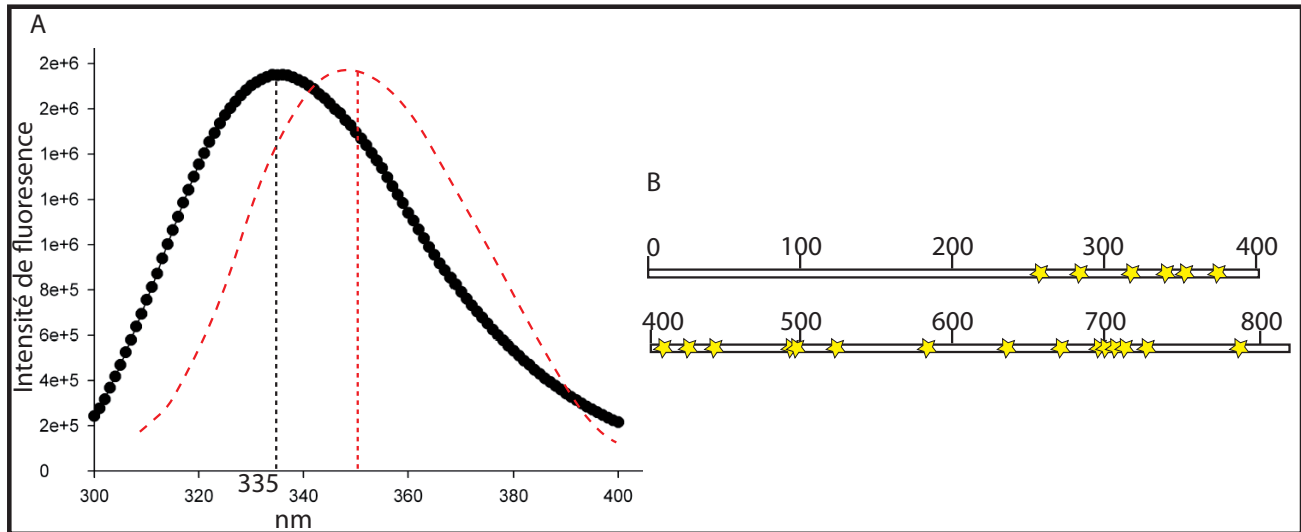


FIG45 : Fluorescence des tryptophanes sur eIF3b

A : Spectre de fluorescence de la protéine eIF3b (noir) à 0,1mM, en rouge spectre théorique de la protéine dépliée.

La longueur d'onde d'excitation est de 288nm et le spectre d'émission est analysé entre 300 et 400 nm.

B : Localisation des 21 Trp sur la protéine eIF3b (jaune).

maximum à 350 nm et qui indiquerait un état déplié (**FIG45A**). Les 3/4 des Trp composant la protéine eIF3b sont répartis dans la partie centrale et C-terminale (**FIG45B**). Ainsi, nous ne disposons pas d'informations sur les 250 premiers acides aminés de la protéine. A ce stade de la caractérisation, nous pouvons constater que la protéine eIF3b sous forme recombinante semble repliée et comporte une majorité de feuilletts β . La protéine est produite sous forme soluble en quantité suffisante, avec un niveau de pureté compatible pour une étude structurale. A l'aide de la plate-forme automatisée de l'EMBL, nous avons donc débuté une campagne de cristallogénèse. La protéine ne pouvant être concentrée à plus de 15mg/ml, nous avons utilisé toute la production pour réaliser 5000 gouttes (3 productions de protéines à 3 concentrations différentes sur 6 plaques contenant 96 conditions de cristallogénèse) qui ont toutes été analysées. Cependant, il n'a pas été possible d'identifier de conditions propices à la cristallisation d'eIF3b. A la suite de ces résultats, nous avons décidé d'utiliser une autre technique pour aborder étudier la structure de cette protéine. Parmi ces techniques, la microscopie électronique n'étant pas adaptée à l'étude de protéine de cette taille, et la RMN ne permet pas l'étude rapide de protéine de taille supérieure à 50kDa, nous avons utilisé la technique de SAXS. Cette technique est parfaitement adaptée pour l'étude de protéines en solution. Possédant déjà une expérience dans l'analyse des données de diffusion (cf Chapitre I des résultats), j'ai pu étudier l'enveloppe de cette protéine dans de bonnes conditions.

III Etude structurale de la protéine eIF3b et son motif MRR par SAXS

III-1 Etude du domaine MRR

Les seules informations structurales disponibles sur la protéine eIF3b correspondent au motif MRR situé dans le domaine N-terminale (185-268) de la protéine (soit 88 résidus sur 814). Afin de produire un modèle de l'enveloppe de la protéine par SAXS le plus précis possible, j'ai produit dans un premier temps le domaine MRR sous forme soluble. Le domaine MRR a été cloné dans le vecteur d'expression pET-TEV entre les sites *NdeI* et *XhoI*. Les conditions de production et d'expression ont été testées et nous indiquent que la protéine est fortement exprimée sous forme soluble et stable à 37°C en milieu LB (**FIG46**). Ayant l'intention d'utiliser les données structurales de ce domaine pour étudier la protéine entière, j'ai procédé à sa caractérisation. Dans un premier temps, nous avons utilisé la technique de MALLS, technique qui utilise la diffusion de lumière multi-angles couplée à la chromatographie d'exclusion de taille et à la réfractométrie (SEC-MALLS-RI). La technique de SEC-MALLS ou Chromatographie d'Exclusion Stérique Multiangle Laser Light Scattering est une méthode de chromatographie en

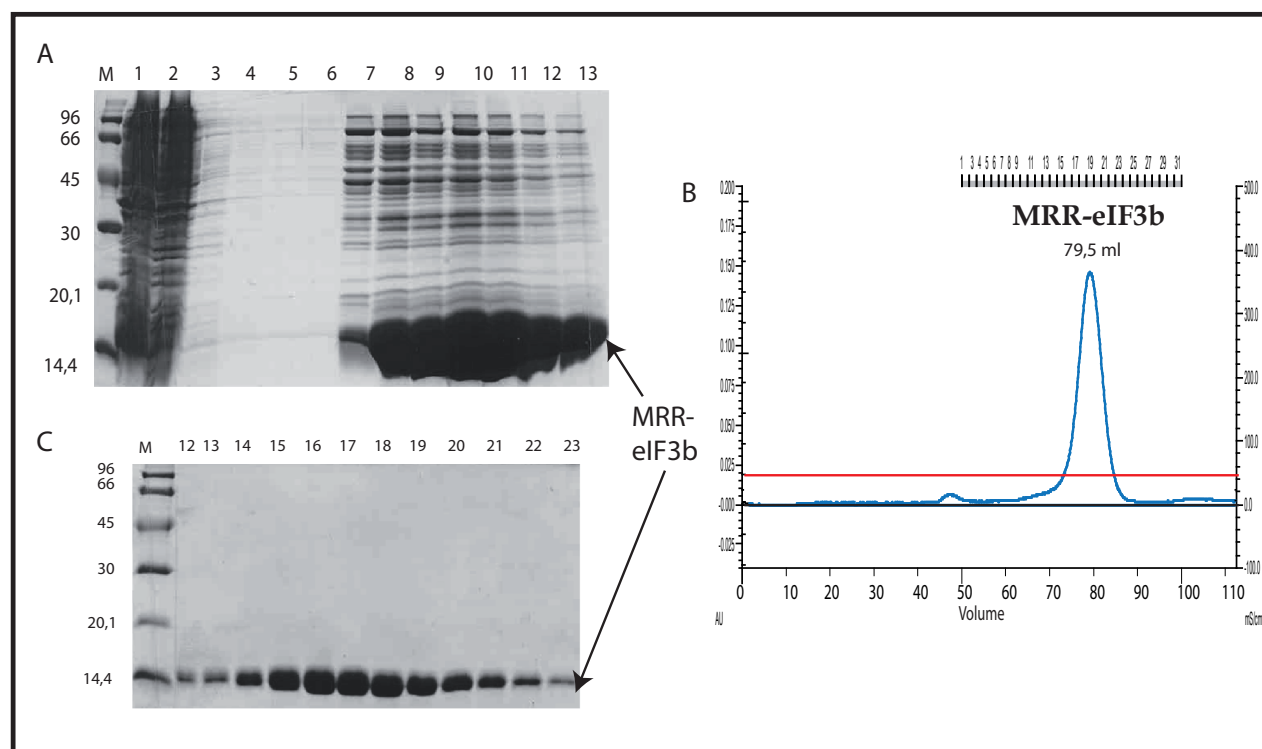


FIG46 : Purification du domaine MRR-eIF3b

A : Gel SDS PAGE 15% après chromatographie d'affinité, M : Marqueur de poids moléculaire.

1 : extrait total. 2 : fractions non retenues 3 - 6 : lavage à 1M NaCl. 7-13 : Fraction d'élution 500mM Imidazole.

B : Chromatogramme obtenu après chromatographie d'exclusion de taille S75 (120ml).

C : Gel SDS PAGE 15% après chromatographie d'exclusion de taille, M : Marqueur de poids moléculaire. 12 à 23 : fractions d'élution.

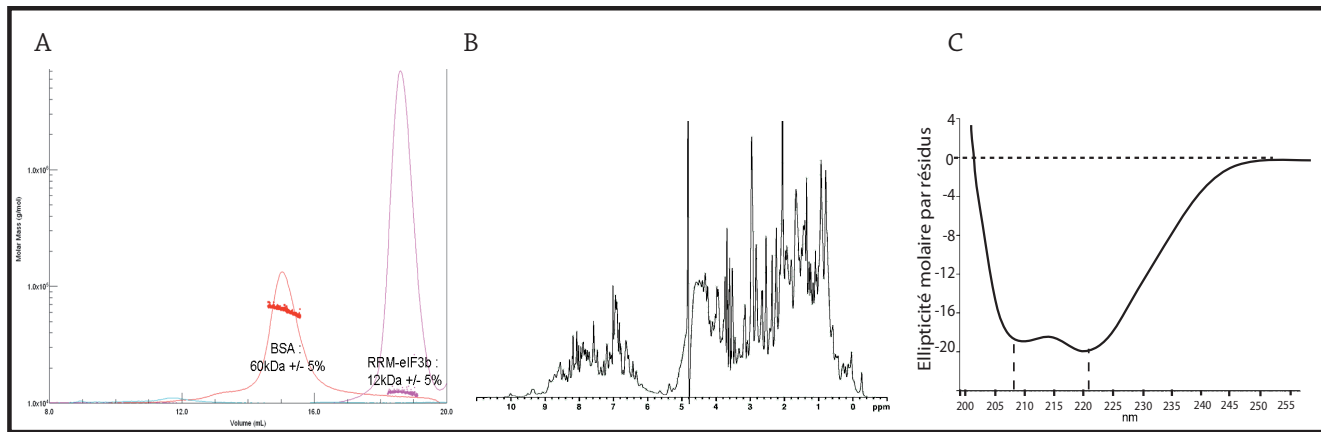


FIG47 : Caractérisation de la protéine MRR-eIF3b

A : Chromatogramme obtenu après analyse de la protéine MRR-eIF3b en MALLS. Les résultats obtenus sont comparés à la protéine standard (BSA de 60 kDa). Le tampon utilisé est celui de la purification par chromatographie de filtration sur gel de la protéine MRR (20mM Tris-HCl pH 7.5; 150 mM NaCl). L'analyse des résultats confirme la présence d'une protéine à la taille attendue sans présence d'oligomérisation.

B : Spectre de RMN 1D (HSQC) de la protéine MRR-eIF3b. La dispersion des pics, correspondant aux protons, est caractéristique d'une protéine structurée.

C : Résultats de dichroïsme circulaire obtenus sur la protéine MRR-eIF3b à 5.10^{-5} M en tampon 20mM tris HCl pH 7,5; 100mM NaF.

phase liquide permettant de séparer et d'identifier des macromolécules en fonction de leur taille. La présence de plusieurs détecteurs (viscosimètre, photomètre, réfractomètre, diffusion de lumière) nous permet de connaître de nombreux paramètres biophysiques des molécules analysées. Cette technique permet aussi d'estimer l'homogénéité des échantillons. La chromatographie d'exclusion de taille est réalisée à l'aide d'une HPLC (chromatographie en phase liquide haute performance). La diffusion de lumière statique est mesurée à 18 angles différents, simultanément à l'aide d'un détecteur DAWN-EOS (Wyatt Technology Corp., Santa Barbara, CA) en utilisant un laser qui émet à 690 nm, l'excès d'indice de réfraction est mesuré à l'aide d'un réfractomètre différentiel RI2000b (Schambeck). Les données obtenues sont analysées avec le logiciel ASTRA V (WyattTechnology Corp.). La figure **47A** montre les résultats obtenus et nous indique la présence d'une protéine homogène et monodispersée de 12kDa possédant une forme globulaire (**FIG10A**). Afin d'étudier l'état de repliement de la protéine MRR-eIF3b, j'ai utilisé plusieurs techniques en parallèle :

- Une analyse en RMN1D auprès du laboratoire LRMN de l'IBS (Institut de Biologie Structurale). La dispersion des pics est caractéristique de la présence d'une protéine structurée (**FIG47B**).
- Une analyse par CD qui indique la présence d'un spectre caractéristique d'une protéine comportant des hélices α (**FIG47C**). La mesure de signal de CD est de -20 soit -4550 en ellipticité molaire, ce qui correspond à 13 % d'hélice α . Ce résultat est compatible avec la prédiction de structure. Cette protéine s'exprime à raison de 30 mg/L de culture (LB), sous forme soluble. Il est possible de la concentrer à 50 mg/ml en présence de 5% de glycérol. Ces résultats nous indiquent que la protéine MRR-eIF3b recombinante est produite de manière soluble, stable et structurée. Au début de ma thèse, la structure de ce domaine n'était pas connue. Disposant de

plusieurs milligrammes de protéine purifié, j'ai procédé à une campagne de cristallogénèse. La protéine a été concentrée jusqu'à 48mg/ml et diluée à 24 et à 12 mg/ml avant d'être utilisée par la plate-forme HTxLab de l'EMBL. Plusieurs conditions ont permis l'obtention de cristaux sous forme de fines aiguilles. Du fait de leurs tailles, j'ai réalisé plusieurs étapes d'optimisation afin d'obtenir des cristaux de taille plus importante (**FIG48A**). Après amélioration des cristaux par ajout de glycérol et augmentation de la concentration en sel, j'ai réalisé une collecte de données à l'ESRF sur la ligne BM14 (**FIG48B**). Cependant, les cristaux se sont révélés très sensibles aux rayons X et nous n'avons pu obtenir qu'une diffraction de la protéine à 9Å. Les conditions de cristallogénèse qui ont permis d'obtenir les cristaux du domaine MRR de la protéine eIF3b sont comme suit :

Condition 1 : 0.2 M sulfate d'ammonium; 20 %w/v PEG 3350, 0,1M acide citrique pH 5

Condition 2 : 0.1 M sulfate d'ammonium, 0,05 M MES ph 5.6; 20 %w/v PEG 8000; 0,01 M chlorure de magnésium

Condition 3 : 4.20% Mme 5000; 0,1 M acide citrique pH 4,2

Ayant l'intention de résoudre cette structure par cristallographie aux rayons X, j'ai produit et purifié plusieurs mg de ce domaine. Cependant au même moment, une équipe concurrente (ElAntak et al. 2007) a publié la structure de ce domaine (présentée dans le paragraphe suivant).

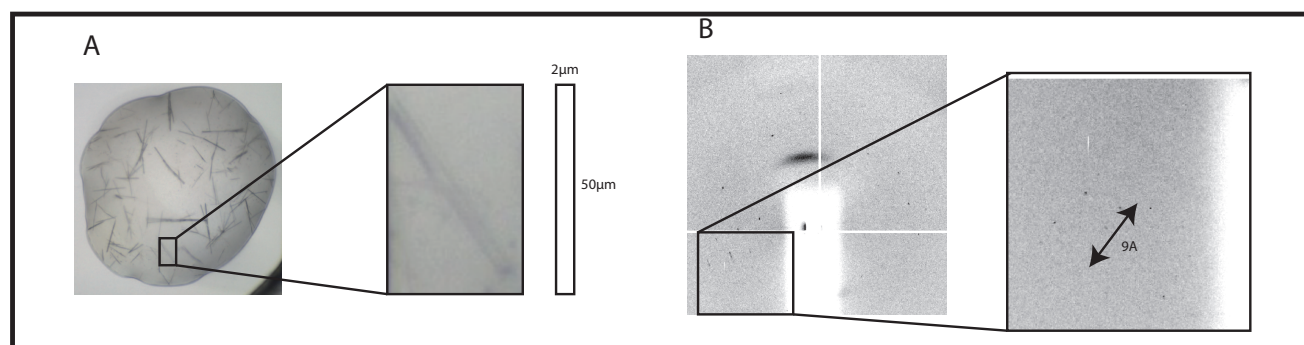


FIG48 : Résultats de cristallogénèse et cristallographie de la protéine MRR-eIF3b

A : Image de cristallogénèse de la condition 2. Les aiguilles obtenues ont une taille de 50μM par 2μM.

B : Cliché de diffraction des aiguilles réalisé à l'ESRF-Grenoble sur BM14. La résolution maximum obtenue n'est que de 9Å.

III-2 Présentation de la structure du domaine MRR.

La structure de ce motif a été résolue par la technique de RMN (ElAntak *et al.* 2007) (**FIG49**). Ce motif présente un feuillet β composé de 4 brins antiparallèles (β_1 : 186-190, β_2 : 217-221, β_3 : 232-237 et β_6 : 262-264) et deux hélices α (α_1 : 200-213, α_2 : 242-249) positionnées perpendiculairement l'une de l'autre soit $\beta\alpha\beta\beta\alpha\beta$. Ce domaine présente également grandes boucles (**L1** 191-199, **L3** : 22-231 et **L5** : 250-256). Ce domaine MRR possède deux RNP (Ribo-Nucleo-Protéin) chargés très négativement RNP1 et RNP2. Ces caractéristiques présentent ce domaine comme un motif MRR non canonique au même titre que le motif MRR-eIF3g. En fonction de ces prédictions, ce motif serait essentiellement dédié à des interactions protéines/protéines. Les auteurs ont pu valider la présence d'une interaction entre le motif MRR-eIF3b et la partie N-terminale de la sous-unité eIF3j (ElAntak *et al.* 2007) et présentent le motif MRR-eIF3b comme impliqué dans une interaction protéine/protéine. Les auteurs ajoutent qu'il n'est pas capable d'interagir avec des nucléotides comme un MRR canonique. Les motifs MRR ont été caractérisés dans les années 1980 (Dreyfuss *et al.* 1988). Ces motifs comportent généralement 80 à 90 acides aminés, ils contiennent une séquence centrale chargée négativement et sont composés d'un minimum de deux séquences RNP. La séquence consensus d'un RNP1 est définie par : L/R-G-F/Y-G/A-F/Y-V/I-x-F/Y (x étant un acide aminé acide). La séquence d'un RNP2 est définie par : I/V/L-F/Y-I/V/L-x-N-L (x étant un acide aminé acide). La fonction première de ces motifs est de présenter une surface chargée, permettant une interaction avec des séquences nucléotidiques ou des protéines. Il est important de noter que le même domaine peut posséder la capacité d'interagir à la fois avec des protéines et des acides nucléiques. Cette dualité n'est pas exclusive et dépend essentiellement de la présence de plusieurs sites RNP dans la protéine.

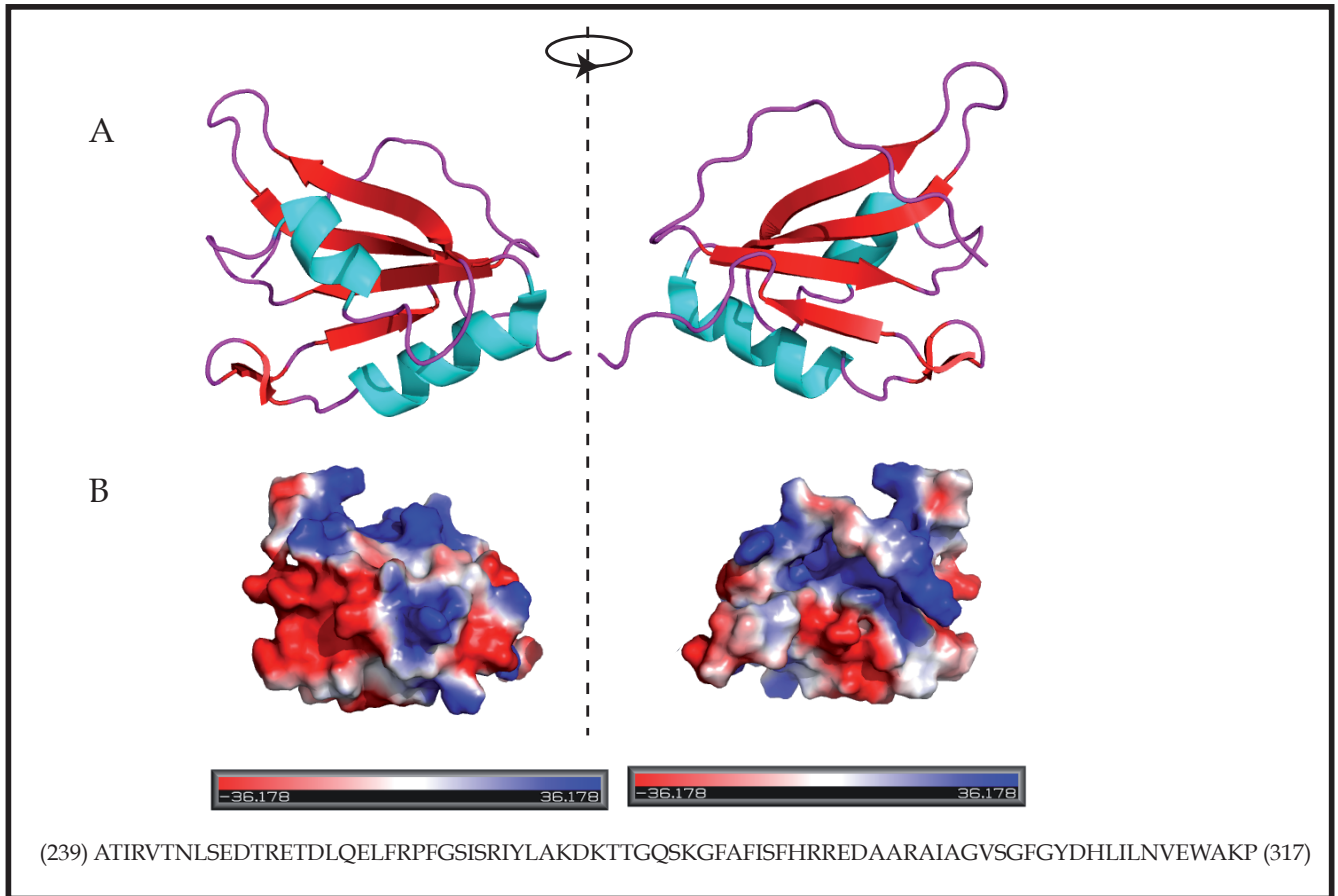


FIG49 : Présentation du motif MRR de la sous-unité eIF3b

A : Représentation en rubant.

B : Représentation en surface électrostatique.

Rouge, zone chargé positivement

Bleue, zone chargé négativement

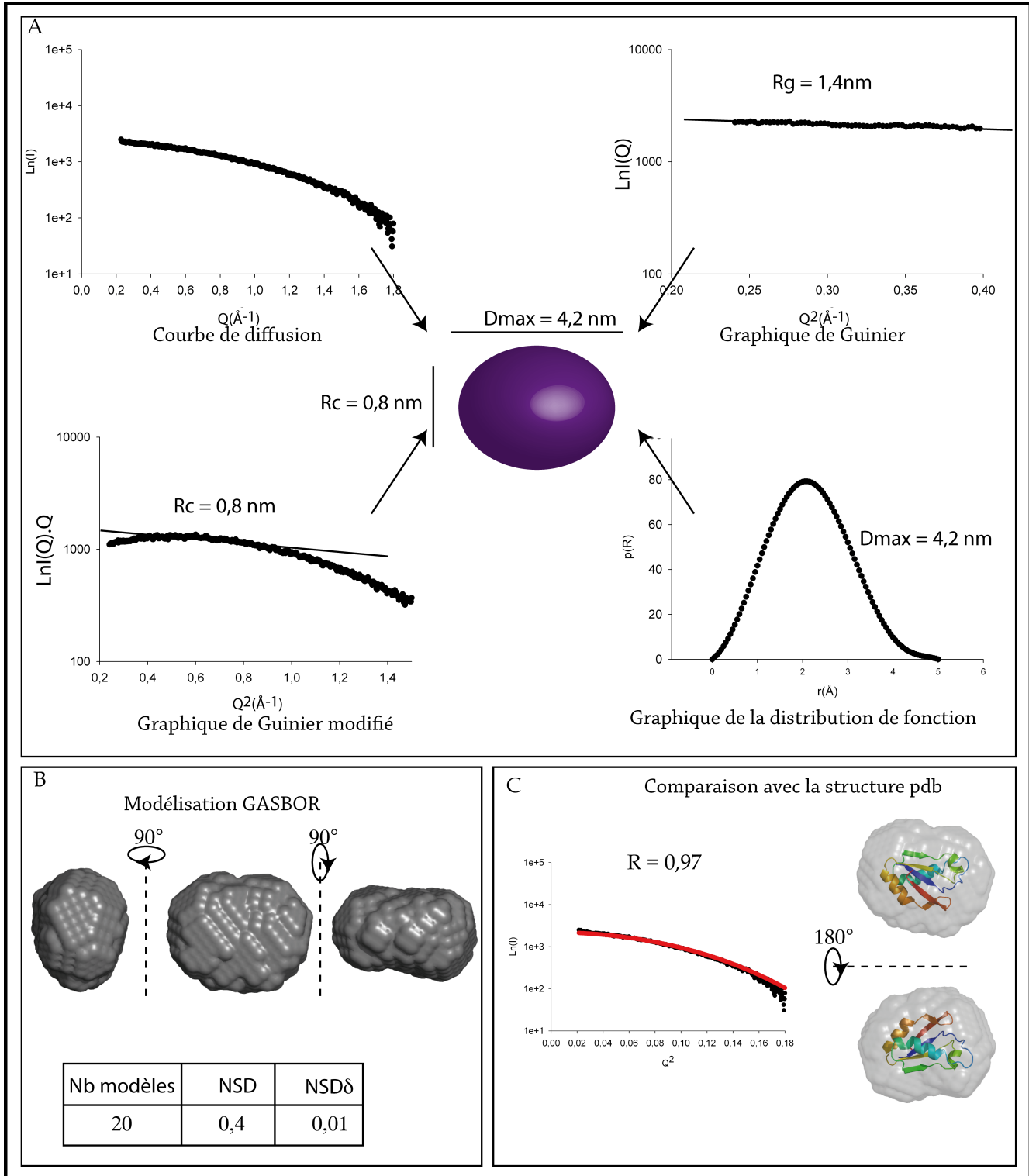


FIG50: Présentation des données biophysiques de la protéine MRR-eIF3b déterminées par SAXS et modélisation de l'enveloppe de la protéine.

A : Traitement des données de diffusion de SAXS et calcul du R_g par le graphique de Guinier, présentation de la $p(R)$ et des données brutes.

B : Modélisation Ab initio de la protéine par la programme GASBOR.

C : Comparaison des données haute résolution (rouge) aux résultats de SAXS (noir).

III-3 Réalisation de l'enveloppe de eIF3b par SAXS

Dans le but d'intégrer la structure du motif MRR dans l'enveloppe de la protéine eIF3b, j'ai tout d'abord généré l'enveloppe du domaine MRR-eIF3b seul en solution. Pour ce faire, j'ai collecté les données de diffusion de lumière de la protéine qui m'ont permis de modéliser, après traitement des données, l'enveloppe de la protéine. Tout d'abord, nous avons déterminé le R_g ainsi que le rayon de section droite (R_c) en utilisant le graphique de Guinier (**FIG50A**). Les données indiquent la présence d'une protéine de forme globulaire faiblement allongée, avec un R_g de 14Å et un R_c de 8Å possédant une taille maximale D_{max} de 42Å. La mesure de distribution de fonction $p(R)$ nous indique la présence d'une protéine de forme globulaire et structurée. Ces données valident les résultats obtenus en RMN, CD et MALLS.

Afin de créer un modèle *Ab initio* de la protéine, j'ai utilisé le programme GASBOR. J'ai réalisé vingt modélisations indépendantes qui ont toutes été compilées à l'aide du programme DAMAVER (**FIG50B**). Les résultats indiquent que la protéine présente une forme de sphère faiblement allongée. Afin de nous assurer de la validité de notre modèle, et disposant de la structure à haute résolution du même domaine, nous avons utilisé le programme SUPCOMB, qui permet de positionner une structure haute résolution dans une enveloppe de SAXS. J'ai pu intégrer la structure haute résolution du domaine MRR dans son enveloppe qui s'intègre au centre de l'enveloppe. Dans un dernier point (**FIG50C**), j'ai comparé la courbe de diffusion expérimentale (noir) à la courbe théorique (rouge) de la protéine. En utilisant le programme CRY SOL, qui permet de générer une courbe de diffusion théorique à partir d'une structure haute résolution, j'ai modéliser la courbe théorique du domaine MRR. Les résultats indiquent que les deux courbes se superposent avec un coefficient de corrélation de 0,97. L'étude conduite sur ce domaine nous permet de valider la méthode de traitement des données et nous permet de disposer de données de référence pour le traitement de la protéine eIF3b. Après avoir obtenu les informations sur le domaine MRR, j'ai procédé de la même façon avec la protéine entière.

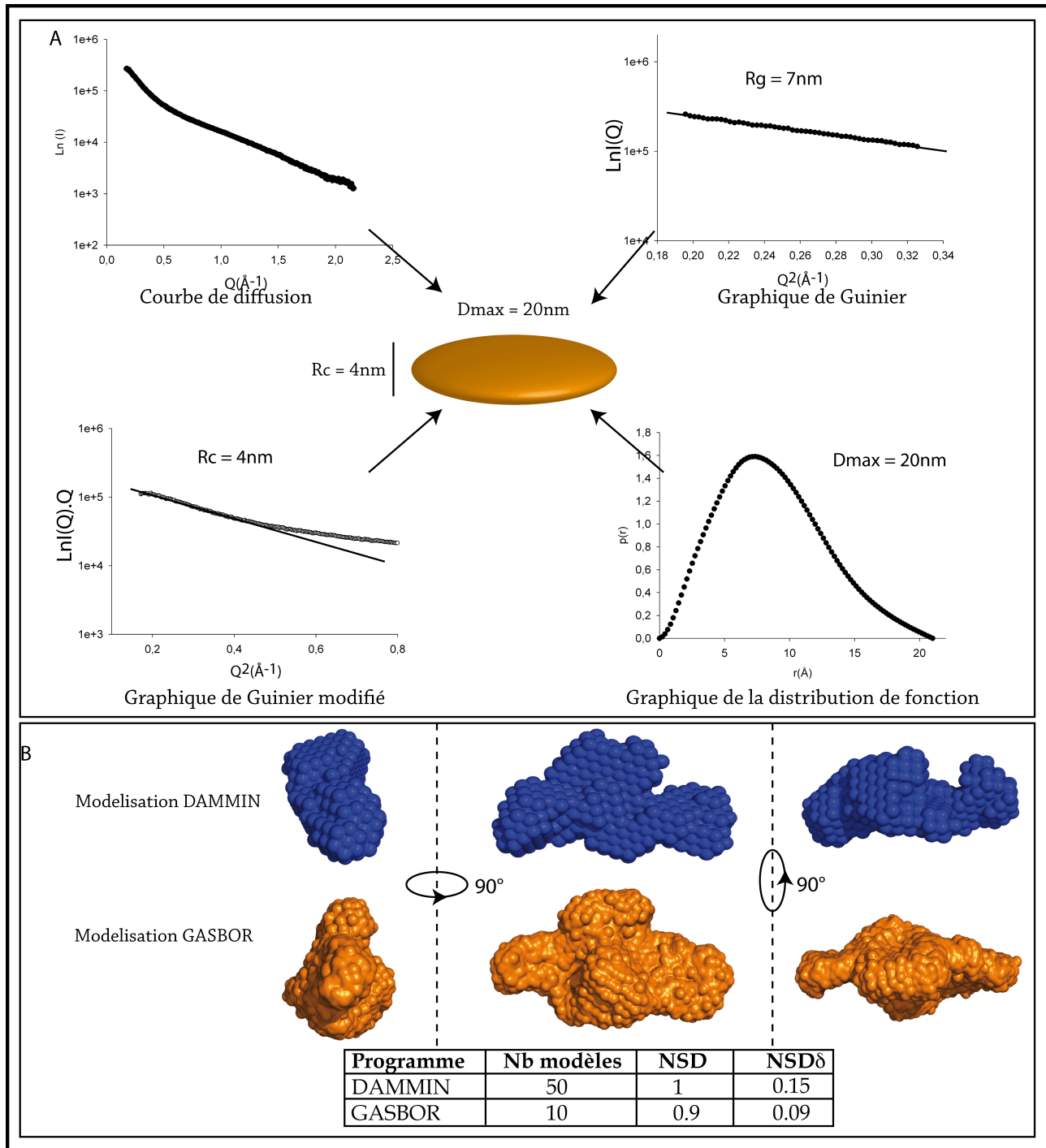


FIG51 : Présentations des données biophysiques de la protéine eIF3b déterminées par SAXS et modélisation de l'enveloppe de la protéine

A : Présentation des informations biophysiques mesurées par SAXS sur la protéine eIF3b. Les informations indiquent la présence d'une molécule de taille sphérique allongée avec une distance maximum de 20 nm.

B : Modélisation *Ab initio* des courbes de diffusion de la protéine eIF3b.

Après avoir collecté les données de diffusion de la protéine eIF3b, j'ai débuté une analyse des paramètres biophysiques avant de procéder à la modélisation de son enveloppe. Tout d'abord, j'ai déterminé le R_g ainsi que le R_c en utilisant le graphique de Guinier (**FIG51A**). Les données indiquent la présence d'une protéine de forme allongée avec un R_g de 70Å et un R_c de 40Å possédant une taille maximal D_{max} de 200Å. La mesure de la distribution de fonction $p(R)$ nous indique la présence d'une protéine allongée. En effet, la courbe s'allonge vers la droite avec un maximum de 200Å. Ce résultat peut s'expliquer par la présence d'une protéine non structurée ou allongée. Ayant validé, par d'autres techniques, le fait que cette protéine était structurée, nous pouvons interpréter la courbe de $p(R)$ en affirmant que la protéine est de forme allongée. Afin de créer un modèle *Ab initio* du domaine MRR, j'ai utilisé les programmes DAMMIN et GASBOR. Pour chaque programme j'ai réalisé cinquante modélisations indépendantes, toutes compilées à l'aide du programme DAMAVER. Les résultats sont présentés dans la **figure 51B**, et nous renseignent sur la forme de la protéine eIF3b. Celle-ci se présente sous la forme d'un disque allongé de 200Å sur 40Å. Le modèle généré par le programme GASBOR a été retenu car il présente un NSDδ le plus faible (tous les modèles générés sont proches les uns des autres). Afin d'optimiser ce modèle en introduisant des données à haute résolution, nous avons utilisé les informations structurales du domaine MRR. Cependant, ce domaine ne représentant que 10% de la protéine et il n'a pas été possible de le localiser au sein de la protéine eIF3b. Ne disposant d'aucune autre information structurale, j'ai initié une étude bioinformatique, afin d'identifier des domaines de la protéine eIF3b dont la structure est connue. Disposant d'informations structurales d'une partie du domaine N-terminal via le motif MRR, je me suis intéressé à la partie C-terminale 350-814. Concrètement, à partir de la séquence primaire de la protéine eIF3b, nous avons effectué une recherche d'homologie de structure par utilisation du serveur LOMETS (LOcal-MEtato-Threading-Server : <http://zhang.bioinformatics.ku.edu/LOMETS/>). Cet outil informatique regroupe plusieurs serveurs de prédictions de structure secondaire (FUGUE, HHsearch, PAINT, PPA-I, PPA-II, PROSPECT2, SAM-T02, SPARKS, SP3). Il utilise neuf programmes pour générer une séquence secondaire consensus. Cette séquence est ensuite utilisée pour rechercher des structures connues de protéines. Si la prédiction de structure est exacte et s'il existe une ou des protéines possédant un domaine homologue, le programme va alors sélectionner ces coordonnées. Ainsi le programme prédit la structure secondaire de cette région et identifie plusieurs structures de protéine. En utilisant différents critères de recherche : comme le pourcentage de la séquence utilisée ou le niveau d'alignement des deux protéines, j'ai pu distinguer une structure répondant à tous les critères. Ces coordonnées correspondent à la protéine WDR5 (Pdb : 2GNQ), qui est impliquée dans la régulation de l'histone H3 (**FIG52**). La fonction de cette protéine n'est pas en relation avec le facteur eIF3, cependant on peut constater que cette protéine forme un anneau composé exclusivement de feuillets β . Cette structure est formée de sept segments arrangés en β -propeller présentant une cavité centrale impliquée dans la liaison d'un petit peptide. Cette solution est la seule qui

a été identifiée par le programme LOMEST avec un score très élevé. Cette structure peut donc correspondre à une partie du domaine C-terminal de la protéine.

Afin de tester cette solution et de tenter de localiser le motif MRR et le modèle prédit par LOMEST dans la protéine eIF3b j'ai utilisé les données structurales identifiées par LOMEST pour modéliser la protéine eIF3b. Pour ce faire, nous avons utilisé le programme de modélisation en rigid body BUNCH (Modelling of Multidomain Proteins Against Multiple Data Sets). Disposant des données de diffusion du domaine MRR de manière individualisé, nous avons pu ajouter ces données dans le processus de modélisation. Après avoir réalisé 20 modélisations de la protéine eIF3b en présence de ces deux structures j'ai vérifié la pertinence des modèles. Il est possible, en analysant les courbes d'ajustement représentant la courbe de diffusion expérimentale de la protéine eIF3b (en noir) et la courbe de diffusion théorique (après intégration des structures à hautes résolutions), d'apprécier la véracité de nos modèles. Comme le montre le graphique de la **figure 53**, il y a une bonne superposition des deux courbes avec un coefficient de corrélation de 0,96. L'interprétation de ces courbes indique que le modèle est compatible avec l'enveloppe de la protéine eIF3b. Grâce à cette approche, nous avons pu localiser le domaine MRR sur un bourgeonnement de la protéine dans sa partie allongée le rendant acces-

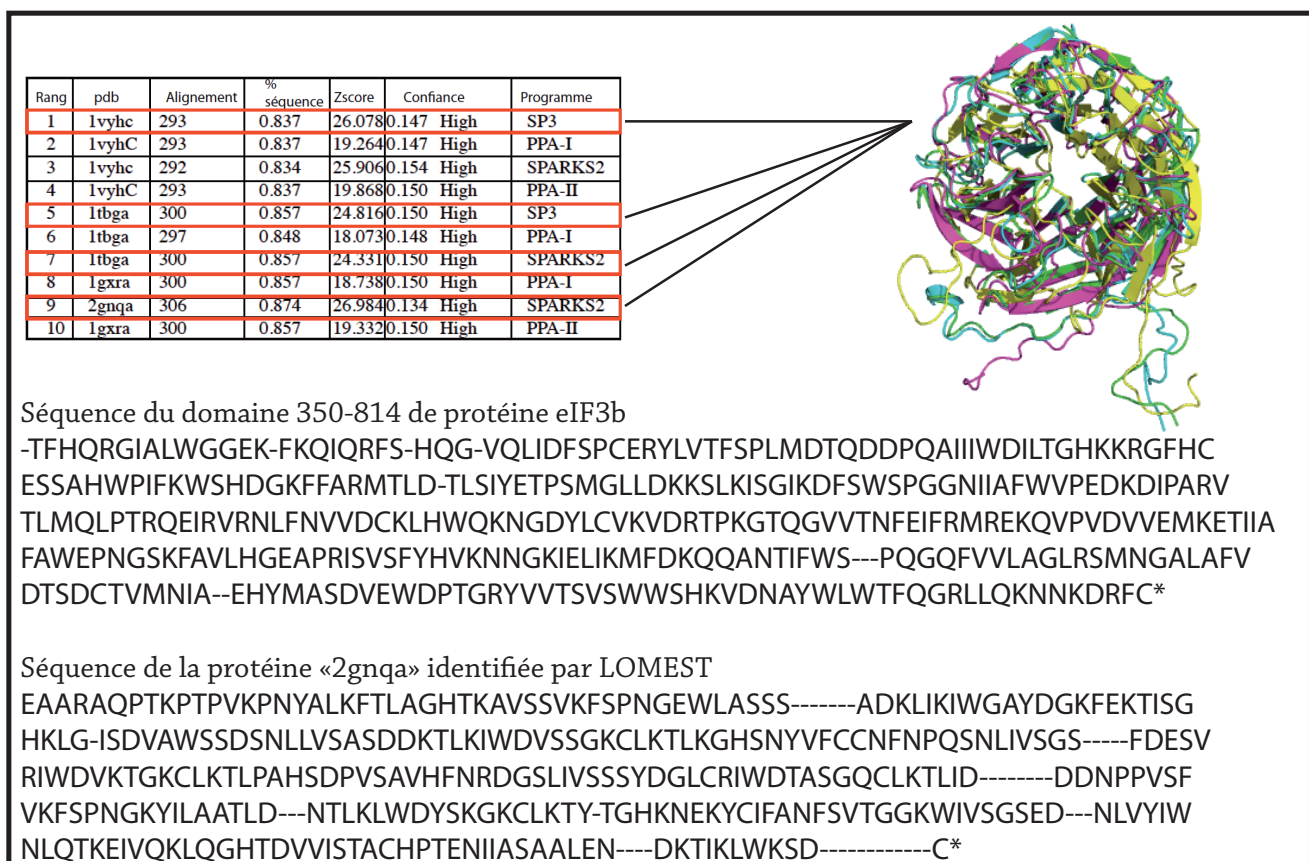


FIG52 : Analyse bioinformatique de la protéine eIF3b par le serveur LOMETS

Résultats de la prédiction de structure par utilisation du serveur LOMETS. La protéine eIF3b a été divisée en 3 sous domaines Nter (1-185), central sans le domaine MRR (270-350) et Cter (350 -814). Le serveur LOMEST a identifié 4 structures ayant une forte homologie avec la prédiction de structure réalisée sur eIF3b. Les résultats obtenus pour le domaine 350-814 indiquent la présence d'un domaine riche en feuillet bêta et formant un anneau de 5nm sur 4 nm.

sible au solvant. Concernant le modèle prédits par LOMEST, celui-ci semble être localisé dans la partie centrale de la protéine. Ne disposant d'aucun autre modèle pour pouvoir comparer nos résultats, nous ne pouvons pas affirmer que ce modèle est valide. Cependant, cette approche de modélisation, qui combine les données de diffusion de SAXS et la prédiction structurale bioinformatiques a déjà précédemment été utilisée pour de petites protéines.

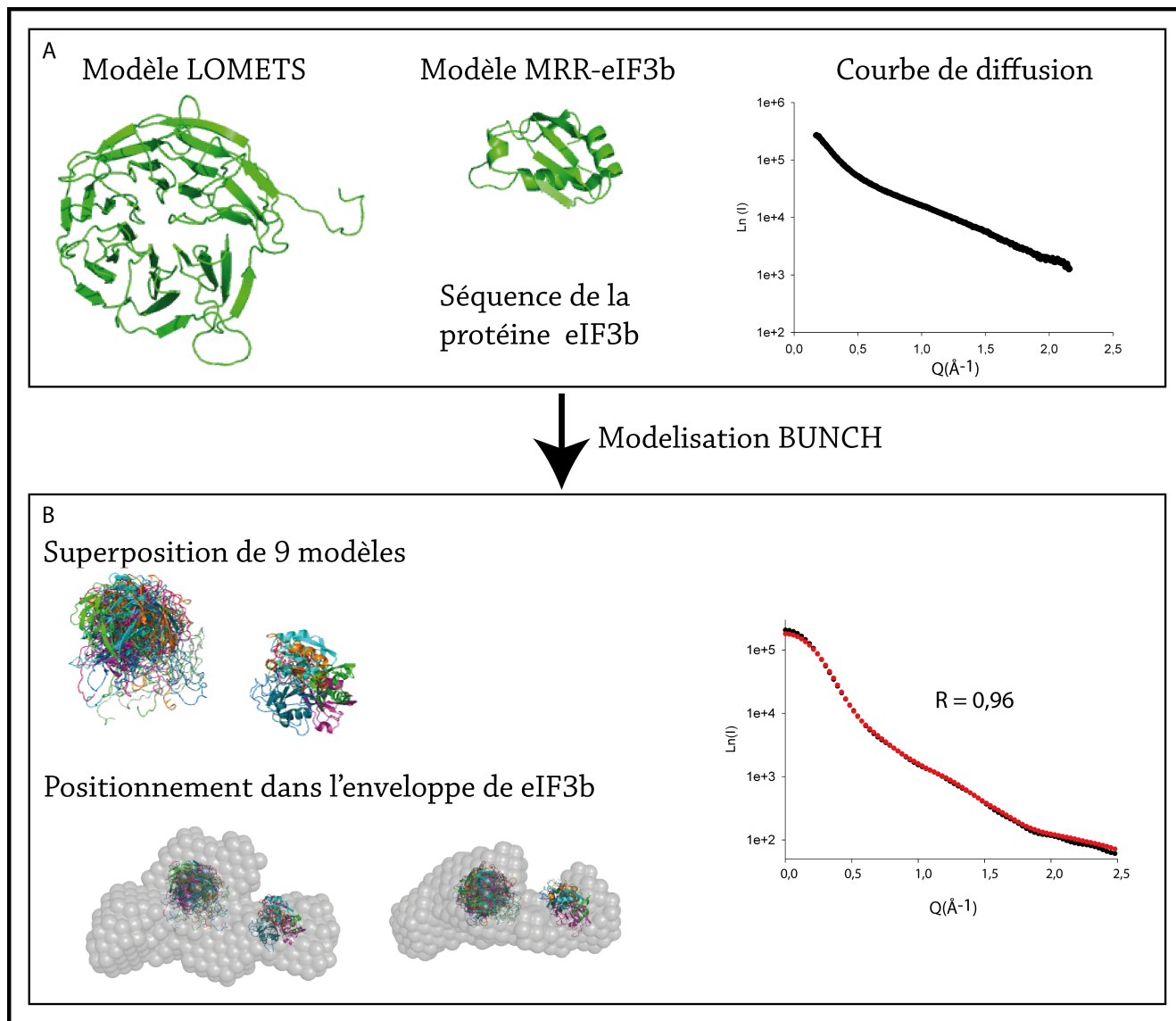


FIG53: Modélisation de la protéine eIF3b par insertion de séquence haute résolution dans l'enveloppe de SAXS.

A : Préparation de la modélisation BUNCH : la structure haute résolution correspondant au modèle LOMETS et au domaine MRR-eIF3b. Séquence primaire de la protéine eIF3b pour le positionnement des structures 3D. Courbes de diffusion de la protéine eIF3b et du domaine MRR.

B : Résultats de la modélisation après utilisation du programme BUNCH. Les résultats de l'ajustement indiquent une bonne superposition de la courbe expérimentale (noire) et de la courbe théorique (rouge). Tous les modèles sont positionnés à l'aide du programme SUPCOMB qui va permettre d'intégrer les structures hautes résolutions dans l'enveloppe correspondant à la protéine eIF3b.

Afin d'utiliser au maximum mes données de diffusions de SAXS, j'ai tenté, dans une dernière approche, de localiser la protéine eIF3b au sein du complexe eIF3. Pour ce faire, j'ai réalisé la mesure de la diffusion du complexe eIF3 par SAXS. Nous avons ensuite utilisé les données de diffusion obtenue pour initier une modélisation à l'aide du programme CREDO (ce programme permet de modéliser les domaines manquants d'un complexe à partir d'une structure). J'ai utilisé l'enveloppe de la protéine eIF3b, présentée plus haut, comme structure de départ pour le programme. Celui-ci a modélisé l'enveloppe du complexe eIF3 (**FIG54**) à partir du modèle de la protéine eIF3b. Cependant les données de diffusion du complexe eIF3 n'étant pas de bonne qualité, je n'ai pas été en mesure de localiser précisément la sous-unité b dans son complexe. Toutefois, en superposant le modèle correspondant à la protéine eIF3b dans l'enveloppe de microscopie électronique du complexe eIF3, nous pouvons constater que la taille et la forme de la sous-unité b est compatible avec le complexe eIF3. Lorsque l'on superpose les enveloppes de eIF3b sur eIF3, on peut se rendre compte que eIF3b occupe une position transversale sur eIF3. Cette disposition peut se montrer favorable pour induire de nombreux contacts avec d'autres sous-unités du complexe. Cependant, il ne s'agit que d'une tentative de localisation et de nombreux contrôles seraient nécessaires avant de pouvoir localiser précisément la sous-unité b dans le complexe eIF3. Les informations structurales détenus sur le facteur eIF3 ne permettent pas d'envisager à l'heure actuelle une telle reconstruction.

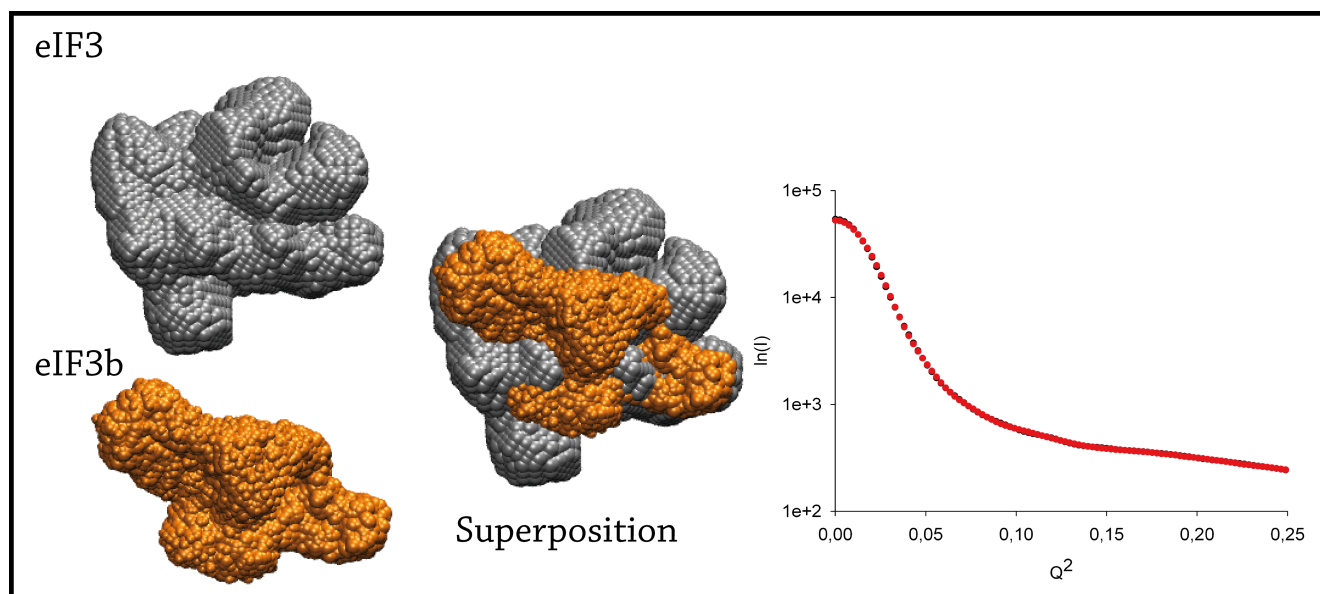


FIG54: Modélisation de la protéine eIF3b dans le complexe eIF3.

L'enveloppe de microscopie électronique correspondant au complexe eIF3 a été transformée en structure pdb contenant des «CA» artificiels représentant l'enveloppe de la molécule. Les données de diffusion du complexe eIF3 ont été utilisées pour positionner la sous-unité b. Concrètement, j'ai utilisé le modèle de la protéine eIF3b comme point de départ de la reconstruction du modèle eIF3 par utilisation du programme CREDO. Le programme a généré une enveloppe intégrant les données de la protéine eIF3b. Après traitement il est possible de localiser la protéine eIF3b au sein du complexe eIF3.

IV Conclusion

Afin de caractériser la protéine eIF3b humaine, nous avons été amenés à produire cette protéine ainsi que son domaine MRR chez la bactérie sous forme recombinante. L'obtention sous forme soluble de la protéine entière comme du domaine MRR en fusion N-terminale avec une étiquette poly-histidine nous a permis d'initier leurs caractérisations par plusieurs techniques biophysiques. Tout d'abord, une analyse par CD a révélée que la protéine eIF3b est majoritairement constituée de feuillets β . Le domaine MRR est quant à lui, composé de plusieurs hélices α ainsi que de quelques feuillets β . Ces premiers résultats nous informent sur la composition en structure secondaire de la protéine. L'interprétation de ces données nous indique la présence d'une protéine structurée composée essentiellement de feuillets β . Au cours de ma thèse, la structure tridimensionnelle du domaine MRR de la sous-unité b du complexe eIF3 a été résolue par une équipe concurrente en complexe avec le domaine N-terminale de la sous-unité eIF3j (ElAntak *et al.* 2007).

A l'issue de ce contre temps, j'ai focalisé mes recherches sur l'étude structurale de la protéine eIF3b sous forme entière. Cette protéine étant de grande taille et ne semblant pas cristallisé j'ai opté pour une étude structurale par SAXS. Après traitement des données de diffusion, j'ai pu modéliser l'enveloppe de la protéine eIF3b et étudier ses paramètres biophysiques. Les données recueillis nous informent sur la taille et la forme de cette protéine. Nous avons pu apprécier la forme globulaire et allongée de la protéine eIF3b qui s'étend sur 200Å de long et qui présente un rayon de 70Å. A l'aide de différents programmes, j'ai pu localiser le domaine MRR ainsi qu'une région riche en feuillets β , préalablement identifiée par le serveur LOMETS dans l'enveloppe de la protéine. La taille et la forme du modèle semblent compatibles avec les données de microscopie électronique qui présentent l'enveloppe du complexe eIF3 comme une protéine polylobée de forme allongée (Siridechadilock *et al.* 2005). En analysant la forme de la protéine eIF3b nous avons pu identifier plusieurs bourgeonnements pouvant correspondre à la partie centrale du complexe eIF3. Ces résultats ne sont que préliminaires et demanderaient à être validés par d'autres études complémentaires. Cependant c'est la première fois que l'enveloppe de la protéine eIF3b est modélisée et il n'existe pas à ce jour de référence structurale pour confronter nos résultats. L'obtention d'informations structurales supplémentaire sur cette protéine nous permettrait de mieux comprendre son agencement au sein du facteur eIF3. C'est dans cette optique que nous avons débuter une campagne de production sous forme recombinante des sous-unité composant le facteur eIF3. Les résultats ne sont que préliminaire, mais il semble possible de produire certaines protéines sous forme solubles et stable (cf **Annexe IV** page 189). Grâce à la technique de SAXS nous avons pu déterminer la forme et la taille de la protéine eIF3b qui est une composante essentielle du complexe eIF3.

La sous-unité b du complexe
eIF3 interagit avec le domaine
IIIId de l'IRES par son motif
MRR

Introduction

L'IRES du VHC possède 342 nt conservés en 5'UTR qui sont repliés en quatre domaines. Ceux-ci sont impliqués dans le recrutement de différents facteurs cellulaires lors de l'initiation de la traduction. Les domaines II, III et IV sont capables d'interagir avec la sous-unité ribosomique 40S. Le domaine apical IIIabc, interagit avec le facteur eIF3 chez l'humain. Ce facteur est composé de 12 sous-unités formant un macrocomplexe de 750kDa (**FIG55A**). Plusieurs études ont été conduites sur la formation du complexe eIF3/IRES. Les données de la littérature nous indiquent que 5 des 12 sous unités (a, b, c, d et f) interagissent avec l'IRES du VHC (Sizova *et al.* 1998) (**FIG55B**). Cependant, l'implication directe de la sous-unité eIF3b dans l'interaction avec l'ARN viral n'a pas été clairement élucidée. Les techniques utilisées pour étudier cette interaction, sont essentiellement basées sur des approches de pontage chimique *in vitro* qui ne permettent pas une étude fine du complexe. De plus, les études conduites jusqu'à présent n'étudiaient que la capacité de l'IRES entier à interagir avec certaines protéines cellulaires. Il est donc difficilement possible de définir une séquence nucléotidique précise impliquée dans une ou plusieurs interactions avec les facteurs cellulaires. Comme cela a été énoncé dans l'introduction, l'IRES du VHC interagit à la fois avec la sous-unité ribosomique 40S et avec le facteur eIF3. Les

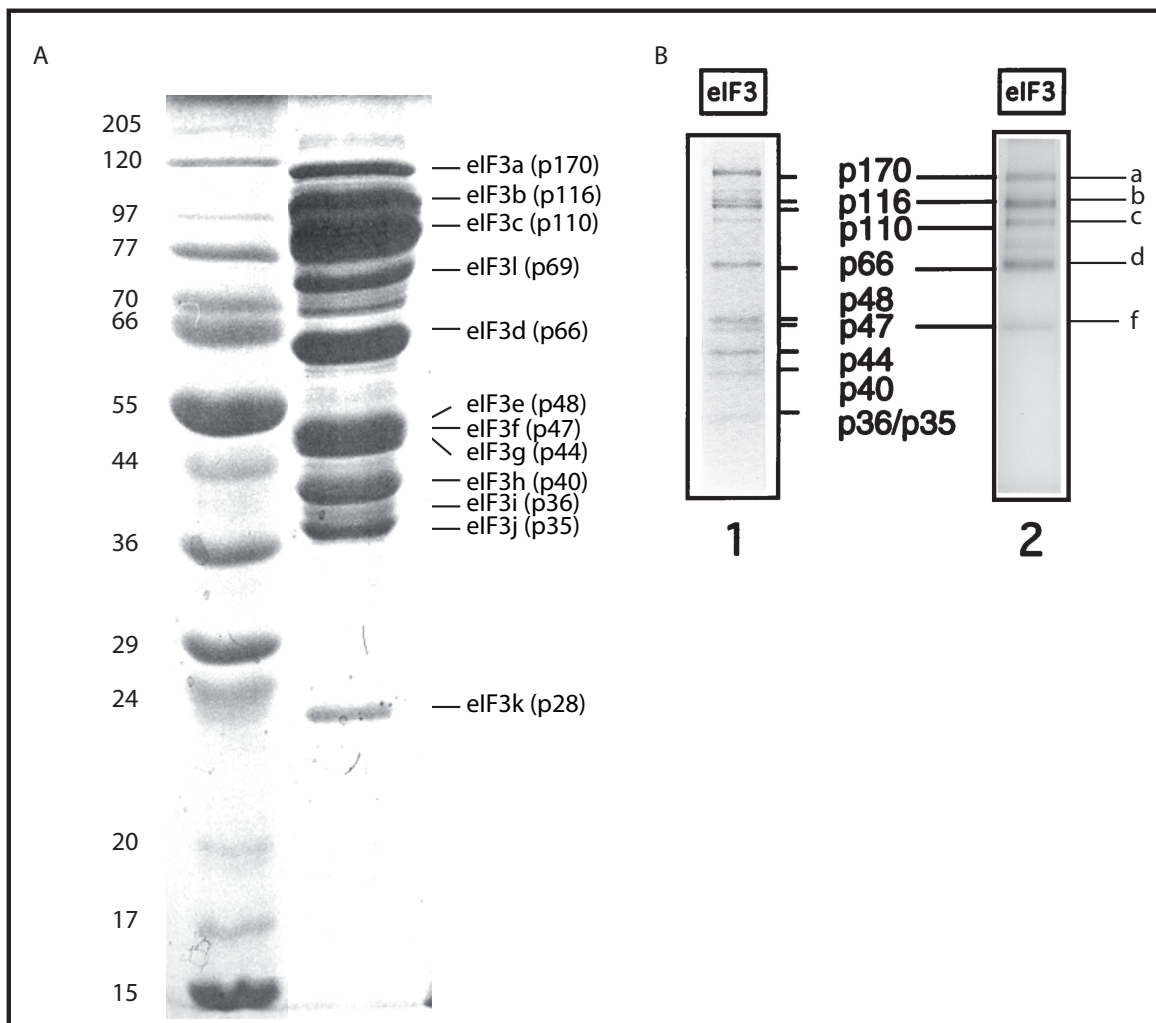


FIG55 : Présentation du complexe eIF3

A : Gel d'acrylamide SDS PAGE 12% du complexe eIF3 (5µg déposés).

B : La piste 1 présente le complexe eIF3 libre et la piste 2 présente les sous-unités de eIF3 détectées après pontage avec de l'ARN de l'IRES du VHC (Sizova *et al.* 1998).

données de la littérature indiquent que l'IRES possède une affinité plus grande pour la sous-unité 40S. Cependant la différence entre l'affinité de 40S (5nM) et eIF3 (30nM) est faible, et il semble difficile de discerner qui de ces deux facteurs se fixe en premier sur l'IRES. Cependant, un modèle existe et suggère que la sous-unité 40S se fixe en premier sur l'IRES (Otto *et al.* 2004). Lorsqu'on s'intéresse au facteur eIF3, on remarque que la protéine eIF3b est impliquée dans de nombreuses interactions protéine/protéine et de ce fait, elle est décrite comme impliquée dans l'échafaudage du complexe eIF3. Cette protéine, qui possède des motifs d'interaction avec les sous-unités eIF3e et eIF3j, possède également un motif MRR dans son domaine N-terminal. Cependant, l'interaction de ce domaine avec l'ARN viral n'a pas encore été établi. La protéine eIF3b est impliquée dans des interactions intra et intermoléculaires lors de l'initiation de la traduction coiffe dépendante et indépendante. En raison de l'importance de cette protéine et de son implication dans divers mécanismes, j'ai voulu étudier si eIF3b était capable d'interagir avec l'ARN viral. Nous avons utilisé plusieurs techniques biochimiques et structurales qui permettent d'étudier la formation de complexe entre une protéine et un acide nucléique. En utilisant la technique de RMN, nous avons pu identifier les acides aminés impliqués dans la formation de ce complexe. Disposant ainsi d'une technique permettant d'étudier la formation d'un complexe avec l'IRES et des protéines nous avons voulu identifier le domaine minimum de l'IRES impliqué dans cette interaction. Après avoir caractérisé l'implication du domaine III dans la formation de ce complexe nous avons procédé à une dissection moléculaire de l'IRES qui nous a permis de tester plus de 30 constructions. A la suite de cette dissection, nous avons, avec succès, identifié une séquence de quelques nucléotides suffisante pour se fixer à eIF3b. Dans un dernier point, nous avons utilisé la technique de SAXS pour modéliser le complexe formé entre l'IRES du VHC et le domaine MRR de la protéine eIF3b.

I La protéine eIF3b interagit-elle avec l'IRES du VHC?

Afin de répondre à cette question, nous avons utilisé les techniques de Filter Binding Assay (FBA) et de retard sur gel (EMSA) qui nous ont permis d'étudier la formation du complexe entre eIF3b et l'IRES. La technique de FBA est un procédé sensible qui permet d'étudier et de mesurer les constantes d'affinités entre un acide nucléique et une protéine. Le principe de cette technique est basé sur les propriétés physiques d'une membrane de nitrocellulose qui retient les protéines. Si l'ARN interagit avec une protéine, on pourra alors détecter la radioactivité sur le filtre qui correspondra à l'ARN radioactif complexé à la protéine. Cette technique a précédemment été utilisée par Kieft (Kieft *et al.* 2001) qui étudiait l'interaction entre 40S et eIF3 envers l'IRES du VHC. Concrètement, 10000 cpm d'IRES sont mis en contact avec des concentrations décroissantes de protéine (10^{-4} à 10^{-9} M). Après incubation, le mélange réactionnel est filtré à travers la membrane de nitrocellulose. Seul l'IRES complexé à la protéine est détectée sur le filtre par la méthode Cerenkov dans un compteur de radioactivité. En utilisant cette technique, nous montrons que la protéine eIF3b interagit avec l'IRES du VHC (**FIG56A**).

Le résultat de cette opération est utilisé pour construire un graphique présentant la concentration de la protéine en fonction de la quantité de radioactivité mesurée et donc du pourcentage d'IRES retenu. L'obtention d'une courbe sigmoïdale nous indique la présence d'une interaction spécifique entre ARN et protéine. Le plateau correspond, au maximum de rétention de l'ARN par la protéine, est exprimé en %. Lorsqu'on utilise une concentration très faible d'ARN (c'est le cas ici), il est possible de calculer le K_d apparent à 50% de rétention de la protéine. Les mesures sont réalisées sur 3 expériences indépendantes avant d'être moyennées. Afin de valider, dans un premier temps, mes conditions expérimentales j'ai utilisé cette technique pour étudier le complexe formé entre l'IRES du VHC et le facteur eIF3 complet. Ce complexe a été purifié à partir de lysat de réticulocyte de lapin. Lors de ma thèse, j'ai initié une collaboration avec l'équipe de Matthias Hentze de l'EMBL Heidelberg qui m'a fait parvenir généreusement de petites quantités de complexe purifié. Grâce à cette collaboration, j'ai pu disposer de suffisamment de protéine pour réaliser différentes expériences d'interaction. Les résultats indiquent que le K_d mesuré entre l'IRES et le facteur eIF3 est de 5nM (**FIG56A**). Ce résultat est cohérent avec la constante de dissociation mesurée par Kieft en 2001 qui était de 35nM. La variation du K_d (facteur 7) entre ces résultats peut s'expliquer par la difficulté rencontrée lors de la détermination de la concentration de complexe eIF3. Ce macro complexe de 750 kDa ne permet pas la

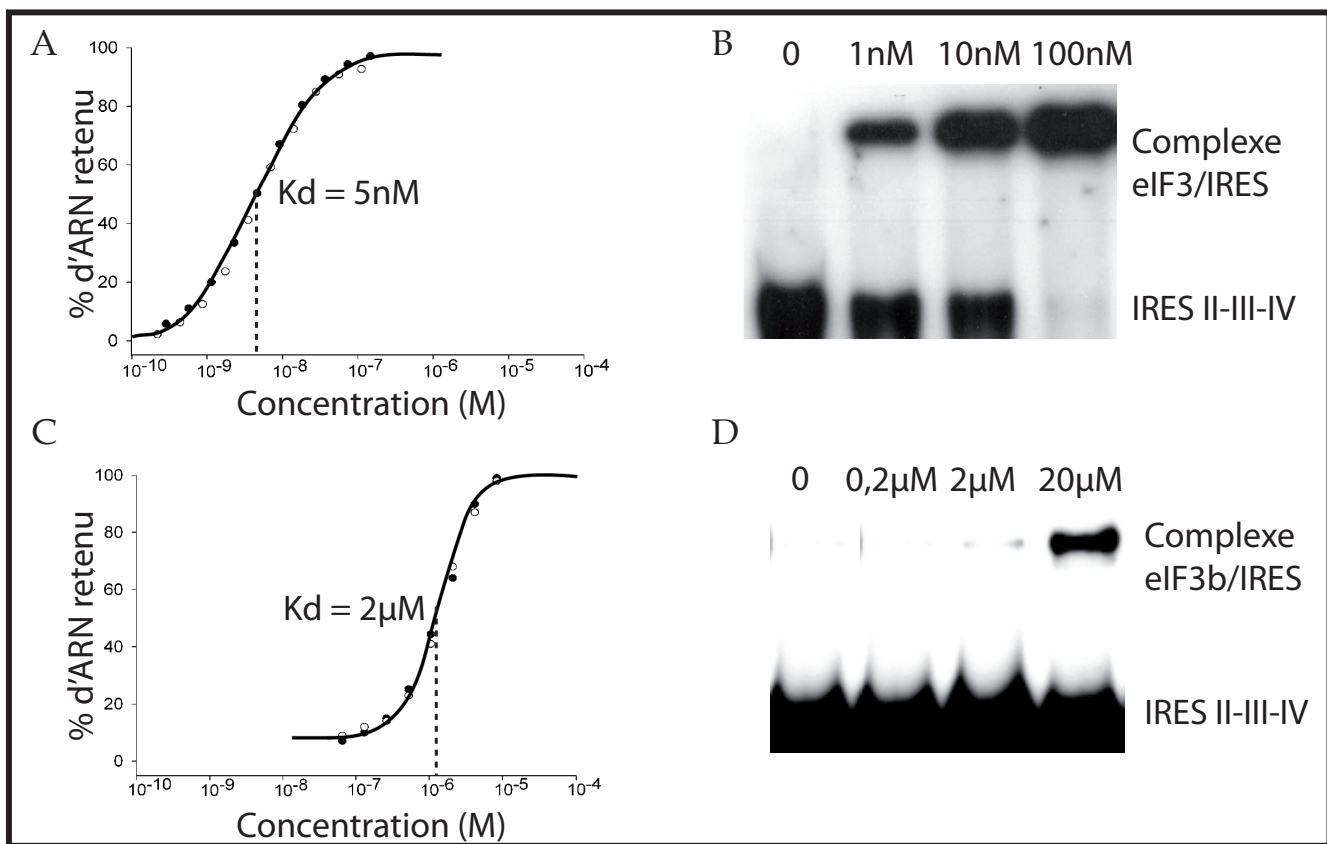


FIG56 : Etude de l'interaction du complexe eIF3 et de la sous-unité eIF3b sur l'IRES du VHC.

A : Courbe expérimentale indiquant l'interaction entre eIF3 et l'IRES.

B : Gel d'acrylamide présentant le retard de la migration de l'IRES en présence du facteur eIF3.

C : Courbe expérimentale indiquant l'interaction entre eIF3b et l'IRES.

D : Gel d'acrylamide présentant le retard de la migration de l'IRES en présence de la protéine eIF3b.

mesure précise de sa concentration par mesure de l'absorbance à 280 nm. Nous avons estimé que la concentration à 1mg/ml correspondait à 1,5 de DO à 280nm. Les faibles variations de la mesure de la concentration en protéine peuvent rapidement avoir un effet sur la mesure du Kd. Afin de confirmer ces résultats j'ai utilisé une seconde technique, qui étudie les variations de migrations entre l'ARN libre ou en complexe avec des protéines. Cette technique de retard sur gel est une approche moins précise que la technique de FBA, cependant elle est rapide et permet de visualiser la présence d'une interaction entre deux molécules par un simple retard de migration. J'ai utilisé les mêmes préparations de protéine et d'ARN pour étudier le retard de migration de l'IRES en présence de concentration croissante de protéine. Comme on peut le voir sur la figure **57B** le complexe eIF3 ralentit la migration de l'IRES : environ 10% de l'IRES est ralenti en présence de 1nM, 50% de l'IRES est ralenti en présence de 10nM et la totalité de l'ARN est ralenti en présence de 100nM du complexe eIF3. En fonction de ces résultats nous pouvons situer le Kd aux alentours de 10nM, confirmant ainsi les mesures de Kd déterminé par la technique de FBA. La protéine eIF3b interagit avec l'IRES avec une constante de 2 μ M, soit 100 à 200 fois moins que l'interaction déterminée avec le facteur eIF3.

Ce résultat amène à deux hypothèses :

Tout d'abord la protéine eIF3b est produite sous forme recombinante, il est donc possible que cette protéine ne possède pas les modifications post traductionnelles requises pour une bonne affinité envers l'IRES.

Il est connu que d'autres sous-unités du complexe eIF3 participent à la l'interaction (a, c, d, f) (Sizova *et al.* 1998). La protéine eIF3b (2 μ M) ne représente que 10 à 15% de l'affinité totale (5nM). Ne disposant pas des autres sous-unités du complexe eIF3 sous forme soluble, je me suis focalisé sur le motif de reconnaissance à l'ARN (185-268) de la protéine eIF3b afin d'établir son rôle dans cette interaction.

II Le domaine MRR de la protéine eIF3b interagit-il avec l'IRES du VHC?

Afin de déterminer si le motif MRR de eIF3b était l'élément responsable de la liaison à l'IRES nous avons entrepris de tester sa capacité à se lier à l'ARN viral. La préparation du domaine MRR-eIF3b a été présentée dans le chapitre II des résultats. Comme témoin de l'expérience, nous avons utilisé le motif MRR de la sous unité eIF3g, qui est impliquée dans une interaction avec de l'ARN ribosomal 18S de la sous-unité 40S (Block *et al.* 1998). Nos résultats indiquent que le domaine MRR-eIF3b se lie bien à l'ARN viral avec un Kd de 2 μ M (**FIG57A et B**). Par contre, le domaine MRR-eIF3g n'interagit pas avec l'IRES (**FIG57C**). Ces résultats nous indiquent que seul le domaine MRR de la protéine eIF3b est responsable de la formation du complexe eIF3b/IRES. En effet, ces deux protéines possèdent la même affinité pour l'IRES. Pour confirmer ces résultats, j'ai testé l'aptitude de la protéine eIF3b sous sa forme entière à entrer en compétition avec son motif MRR lors d'une interaction avec l'IRES. Afin de mener à

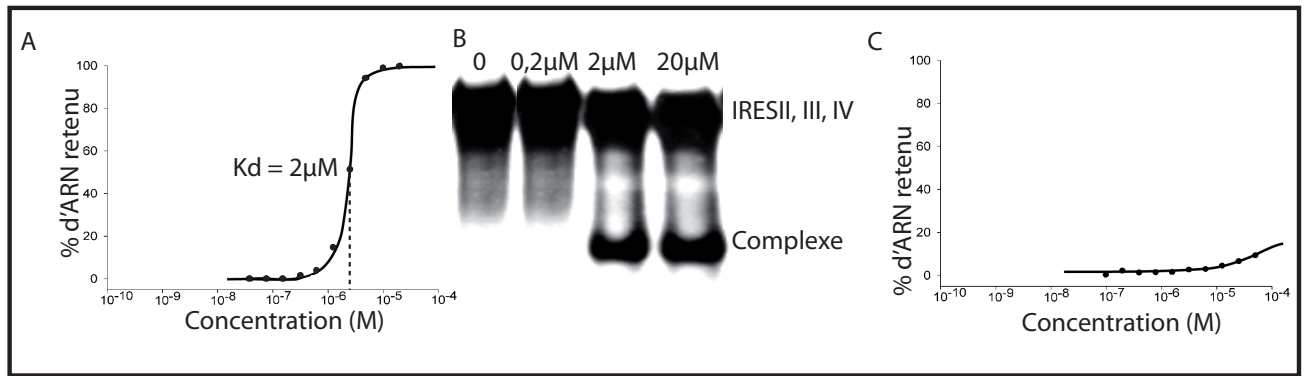


FIG57 : Etude de l'interaction des domaines MRR avec l'IRES du VHC par FBA et EMSA.

A : Courbe expérimentale de l'interaction entre MRR-eIF3b et l'IRES.

B : Gel d'acrylamide présentant le retard de la migration de l'IRES en présence du domaine MRR-eIF3b.

C : Courbe expérimentale de l'interaction entre MRR-eIF3g et l'IRES.

bien cette expérience, il est nécessaire de s'assurer de la réversibilité du complexe MRR-eIF3b/IRES. En effet si le complexe formé, entre la protéine et l'ARN viral, n'est pas réversible il ne sera pas possible d'étudier de compétition. Pour ce faire, j'ai formé le complexe entre l'IRES et le domaine MRR. Après incubation j'ai ajouté des concentrations croissantes d'IRES non marqué et j'ai analysé le % d'ARN. Les résultats indiquent la présence d'un re-largage de l'IRES marqué en présence d'un excès d'IRES. Les résultats sont reportés sur un graphique et présentent le pourcentage d'ARN retenu sur le filtre en fonction de la concentration en ARN. Cet IRES en excès prend la place de l'ARN radioactif avec une constante d'association K_a qui est identique au K_d de $2\mu\text{M}$ (**FIG58A**). Ce résultat nous indique que le complexe est dynamique et qu'il est possible d'en déplacer l'équilibre en présence d'un excès d'un des partenaires. Nous avons ensuite testé la capacité de la protéine eIF3b ou d'eIF3 à lier l'IRES et à entrer en compétition avec le MRR. Cette expérience a pour but de valider la présence d'un seul et unique site de fixation de l'IRES sur la protéine eIF3b. J'ai tout d'abord réalisé le complexe entre l'ARN marqué et la protéine MRR-eIF3b en suivant le protocole standard. J'ai ensuite ajouté des concentrations croissantes de protéine (eIF3b ou eIF3). Les résultats présentés dans la **figure 58B** nous indiquent que la protéine eIF3b n'entre pas en compétition avec son motif MRR car on retrouve une constante K_a de $2\mu\text{M}$ qui est identique au K_d mesuré avec le MRR sur l'IRES. En revanche, on constate que le facteur eIF3 possède un K_a de 10nM et qu'il lie l'IRES du complexe MRR/IRES. Ces résultats montrent bien la présence d'autres motifs de reconnaissance de l'IRES du VHC sur le complexe eIF3. Après avoir confirmé l'implication du domaine MRR dans la formation d'un complexe eIF3, nous avons tenté de perturber la formation de ce complexe par l'action d'un anticorps ou d'un peptide mimétique.

Je disposais, au laboratoire, d'anticorps monoclonaux spécifiques des domaines MRR des protéines eIF3b et eIF3g. Pour tester la capacité de ces anticorps à perturber la formation du complexe MRR/IRES, j'ai utilisé ces anticorps. Les résultats sont présentés dans la **figure 58C** et indique que la présence d'un anticorps spécifique du domaine MRR permet de diminuer de 25% la formation du complexe à une concentration de $1\mu\text{M}$. Comme contrôle, nous avons uti-

lisé un anticorps dirigé contre l'étiquette poly-histidine de la protéine ainsi qu'un anticorps anti MRR-eIF3g. Les résultats indiquent que seule la présence de l'anticorps anti MRR conduit à une diminution de la formation du complexe. Cet anticorps doit donc masquer une partie du domaine MRR impliqué dans la formation du complexe. Ne disposant pas de suffisamment d'anticorps pour réaliser l'analyse de la séquence de l'épitope et d'identifier la séquence peptidique impliquée je me suis rapproché de la séquence primaire du motif MRR pour tester l'implication des deux motifs RNP composant le domaine MRR. En me basant sur l'hypothèse que le motif RNP1 est impliqué dans des interactions protéine/protéine et que le motif RNP2 est dédié aux interactions nucléotides/protéine (Maris *et al.* 2005) j'ai fait synthétiser deux peptides correspondant aux séquences RNP1 et 2 du domaine MRR-eIF3b (RNP1 : KGYIFLGY et RNP2 : VIVVDNV). En utilisant la technique de FBA j'ai testé la capacité de ces peptides à entrer en compétition avec le domaine MRR-eIF3b. Les résultats sont présentés dans la **figure 58D** et indiquent que le peptide RNP2 utilisé à 4 μ M est capable de perturber (50%) la formation du complexe entre le MRR et l'IRES. Ce résultat indique que la séquence du RNP2 est impliquée dans la reconnaissance de l'ARN. Afin de combiner ces résultats, j'ai testé la capacité de l'anti-

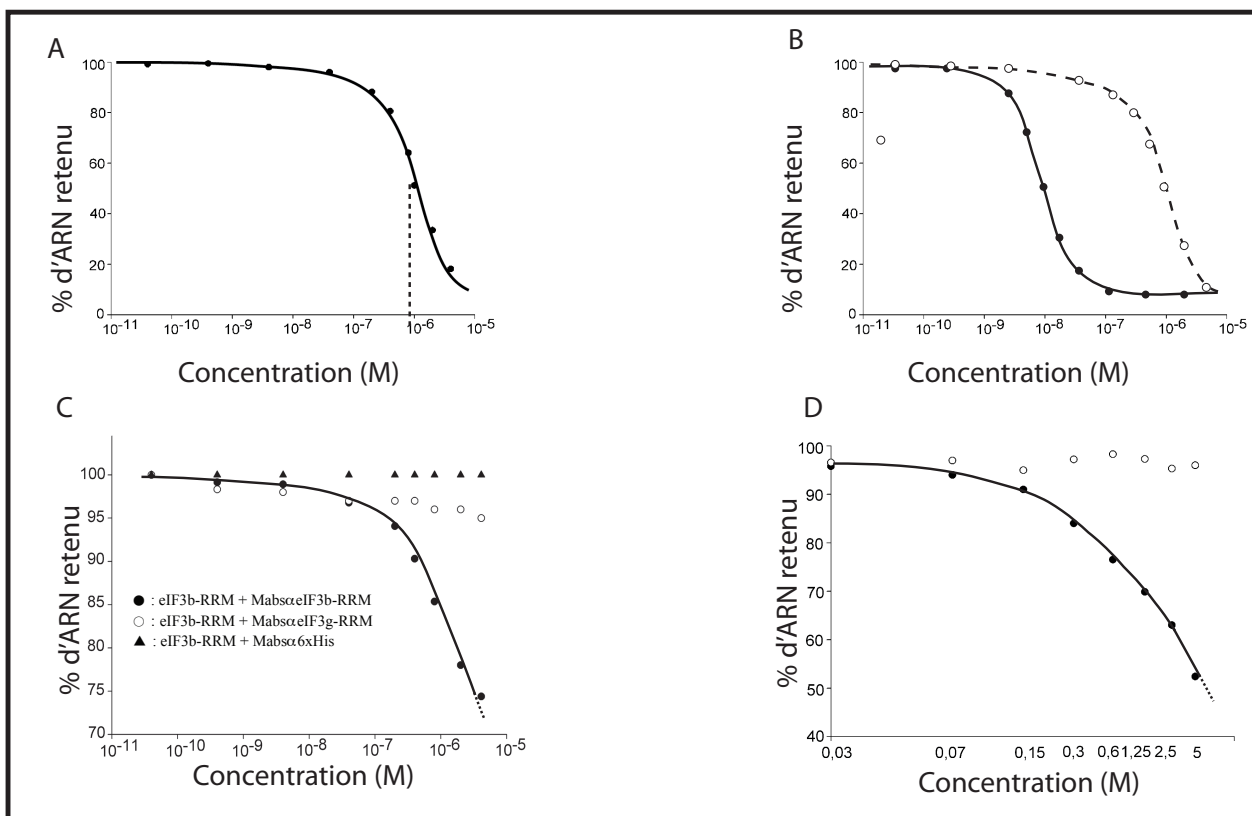


FIG58 : Etude du complexe MRR-eIF3b / IRES, test de compétition et inhibition

A : Le complexe MRR-eIF3b / IRES 32P est en compétition avec de l'IRES II, III, IV froid en excès.

B : Le complexe MRR-eIF3b / IRES 32P est en compétition avec différentes protéines.

Cercle blanc : concentration croissante de protéine eIF3b.

Cercle noir : concentration croissante de eIF3.

C : Un premier complexe protéine anticorps est formé est incubé 30 minutes à 37°C. Ce complexe immun est ensuite utilisé pour former le complexe protéine / IRES dans les conditions standards.

D : Le complexe MRR-eIF3b / IRES 32P est en compétition avec différents peptides.

Cercle blanc : concentration croissante du peptide 1 (KGYIFLGY).

Cercle noir : concentration croissante du peptide 2 (VIVVDNV).

corps anti MRR à reconnaître le peptide RNP2. Cependant, l'anticorps ne semble pas capable de reconnaître ce petit peptide. L'interprétation de ces résultats nous permet de proposer l'hypothèse suivante : Soit l'anticorps reconnaît une autre séquence du domaine MRR qui est elle aussi impliquée dans l'interaction avec l'IRES, soit l'anticorps reconnaît une structure qui n'est pas reformée par le peptide. Afin de déterminer avec exactitude la séquence d'acide aminé impliquée dans la formation de ce complexe, nous avons utilisé la technique de RMN.

III Quels sont les acides aminés du domaine MRR-eIF3b impliqués ?

En collaboration avec le laboratoire de RMN (LRMN) de l'Institut de Biologie Structurale (IBS), j'ai initié une étude de spectroscopie RMN. Nous avons produit dans un premier temps de la protéine MRR-eIF3b sous forme doublement marquée $^{15}\text{N}/^{13}\text{C}$ afin de réaliser l'attribution des pics et d'identifier chaque acide aminé. Nous avons ensuite formé le complexe entre l'IRES non marqué et le domaine MRR-eIF3b marqué isotopiquement (^{15}N). Afin d'étudier la formation du complexe, nous avons effectué la mesure du spectre en 2D à deux concentrations d'ARN (protéine : ARN => 1.2:1 et 4:1). Les détails concernant les expériences sont présentés dans la partie matériels et méthodes, ainsi que dans la publication (Perard *et al.* 2009).

Quand on compare les déplacements chimiques des résidus de la protéine seule ou en complexe avec l'IRES, nous pouvons identifier 15 résidus situés dans deux régions qui sont impliqués dans la formation du complexe. La **figure 59A** illustre trois déplacements chimiques concernant les résidus V192, D256 et Q258. Après analyse de tous les déplacements chimiques, qui sont représenté graphiquement dans la **figure 59B**, nous pouvons discerner l'implication de deux régions localisée entre les aa N191/G196 et les aa K254/T260 du domaine MRR. L'analyse des déplacement chimiques nous indique la présence de deux types de résidus, ceux fortement impliqués (qui possède un déplacement chimique supérieur à 0,1) et ceux moyennement impliqués (qui sont compris entre 0,05 et 0,1). En utilisant cette discrimination, il est possible d'identifier 6 résidus qui sont plus fortement impliqués : Q194, K254, D256, Q258, H259 et T260. On peut noter que 5 de ces résidus sont localisés entre les feuillets β_4 et β_5 faisant de cette région la plus fortement impliquée. Le motif RNP2 fait participer les aa I187-N191-V192-Q194 dans la liaison. La **figure 59C** représente les acides aminés impliqués dans l'interaction. En rouge ceux qui sont fortement impliqués et en orange ceux qui sont moyennement impliqués. L'analyse de la surface électrostatique du motif MRR nous permet de localiser tous ces acides aminés sur une zone située entre une région chargée positivement et négativement. Ce mode d'interaction est différent des MRR canoniques définis jusqu'à ce jour et ne semble pas faire intervenir directement les motifs RNP2 et RNP1.

Après avoir identifié les acides aminés du domaine MRR impliqués dans l'interaction avec l'IRES du VHC, nous avons recherché le motif minimum de l'IRES impliqué dans la formation du complexe.

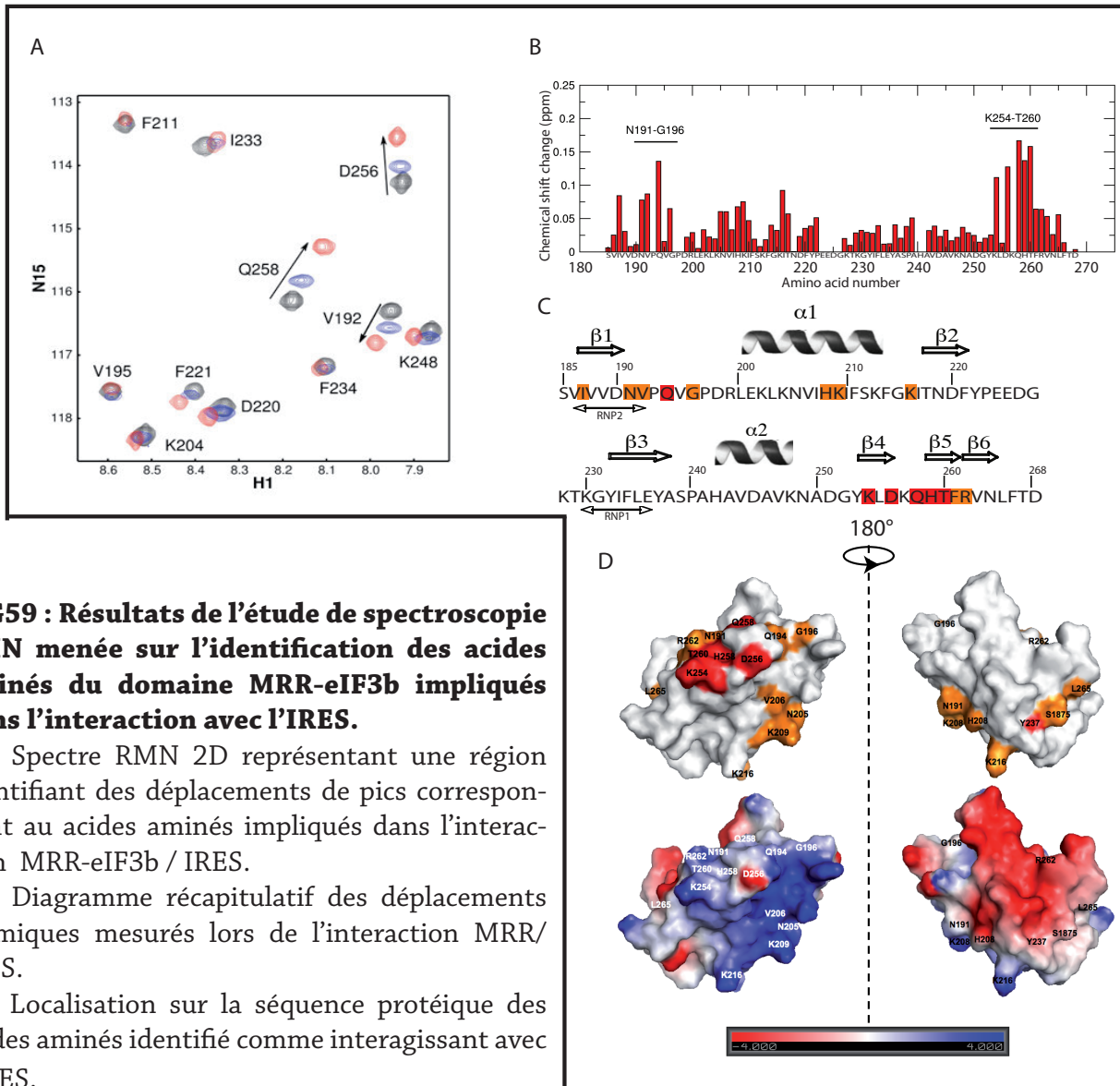


FIG59 : Résultats de l'étude de spectroscopie RMN menée sur l'identification des acides aminés du domaine MRR-eIF3b impliqués dans l'interaction avec l'IRES.

A : Spectre RMN 2D représentant une région identifiant des déplacements de pics correspondant au acides aminés impliqués dans l'interaction MRR-eIF3b / IRES.

B : Diagramme récapitulatif des déplacements chimiques mesurés lors de l'interaction MRR/IRES.

C : Localisation sur la séquence protéique des acides aminés identifié comme interagissant avec l'IRES.

D : Représentation de l'enveloppe du domaine MRR avec identification de la région impliquée dans l'interaction avec l'IRES du VHC. La représentation des surfaces électrostatiques nous indique que cette région est située dans une zone peu chargée.

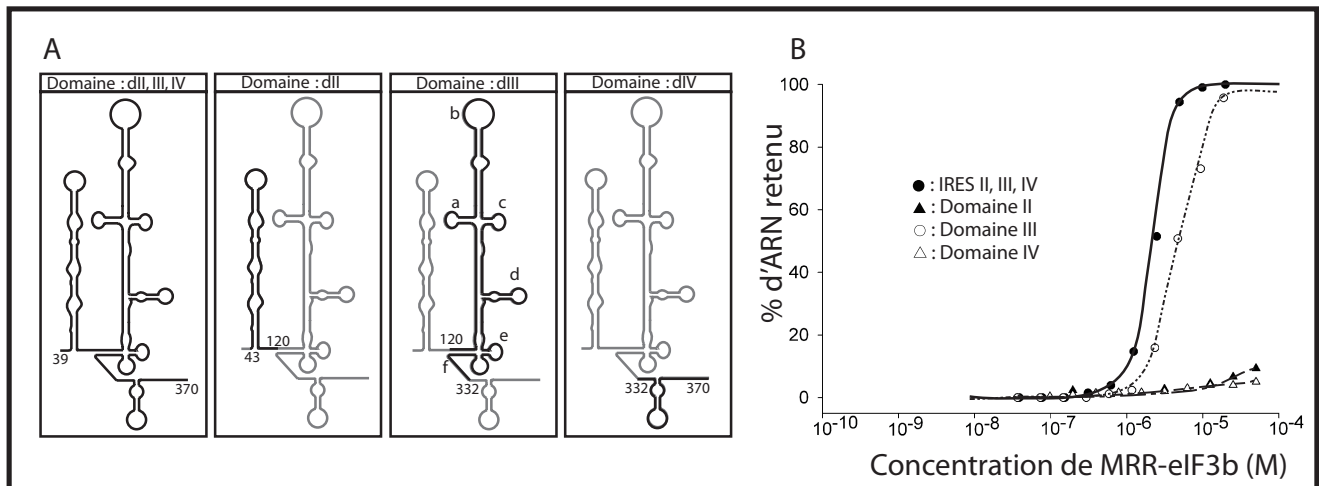


FIG60 : Identification du rôle du domaine III de l'IRES dans l'interaction avec le domaine MRR-eIF3b.

A : Dissection moléculaire de l'IRES et isolement des domaines II, III et IV.

B : Courbe expérimentale indiquant l'interaction entre le MRR-eIF3b et les domaines II, III, IV et IRES entier.

IV Identification du domaine minimum de l'IRES du VHC impliqué dans l'interaction

Dans le but de localiser le domaine minimum de l'IRES impliqué, j'ai tout d'abord testé la capacité de chaque domaine à se lier au domaine MRR-eIF3b par la technique de FBA. Les résultats nous indiquent que seul le domaine III est impliqué dans la formation du complexe et présente un Kd de 5 μ M (**FIG60**).

Afin d'étudier plus en détail l'interaction de l'IRES et de eIF3b-MRR j'ai construit plusieurs mutants de cet ARN par dissection moléculaire (**FIG61**). Une partie de ces mutants a déjà été présentée dans la première partie des résultats. Le **tableau XIV** présente les caractéristiques biochimiques de ces constructions ainsi que les constantes de dissociation mesurée avec le domaine MRR-eIF3b.

Tableau XIV : Présentation des domaines de l'IRES générés par biologie moléculaire et de leur affinité pour MRR-eIF3b.

Nom	Longueur en nt	Position	Mesure du Kd μ M	Présence du domaine IIIId
II-III-IV	343	39-372	2	oui
II-III	291	39-330	4	oui
III-IV	236	134-370	5	oui
III	196	134-330	5	oui
II-IIIabcd	251	39-290	6	oui
IIIdefIV	140	120-142/252-370	10	oui
IIIabcd	156	134-290	10	oui
IIIId	30	253-279	20	oui
II	77	43-120	> 30	non
II-IIIabc	213	39-252	> 30	non
IIIefIV	120	250-370	> 30	non
IV	40	330-370	> 30	non
IIIabc	111	141-252	> 30	non
IIIb1	55	172-227	> 30	non
IIIb2	37	181-218	> 30	non
IIIef	94	120-134/290-370	> 30	non

Les résultats présentés dans le tableau ci-dessus indiquent que les domaines II-III, III-IV, III, III-IIabcd, IIIdefIV, IIIabcd et IIIId possèdent une affinité supérieure à 30 μ M. Tous les domaines possédant le fragment IIIId ont une affinité nettement meilleure pour l'IRES. Par exemple, le fragment IIIabcd possède une affinité de 10 μ M, si l'on supprime le domaine IIIId, on se retrouve à une affinité de 300 μ M. La présence du domaine II, bien qu'il ne soit pas directement impliqué dans l'interaction avec la protéine semble augmenter l'affinité. Cependant, d'après nos résul-

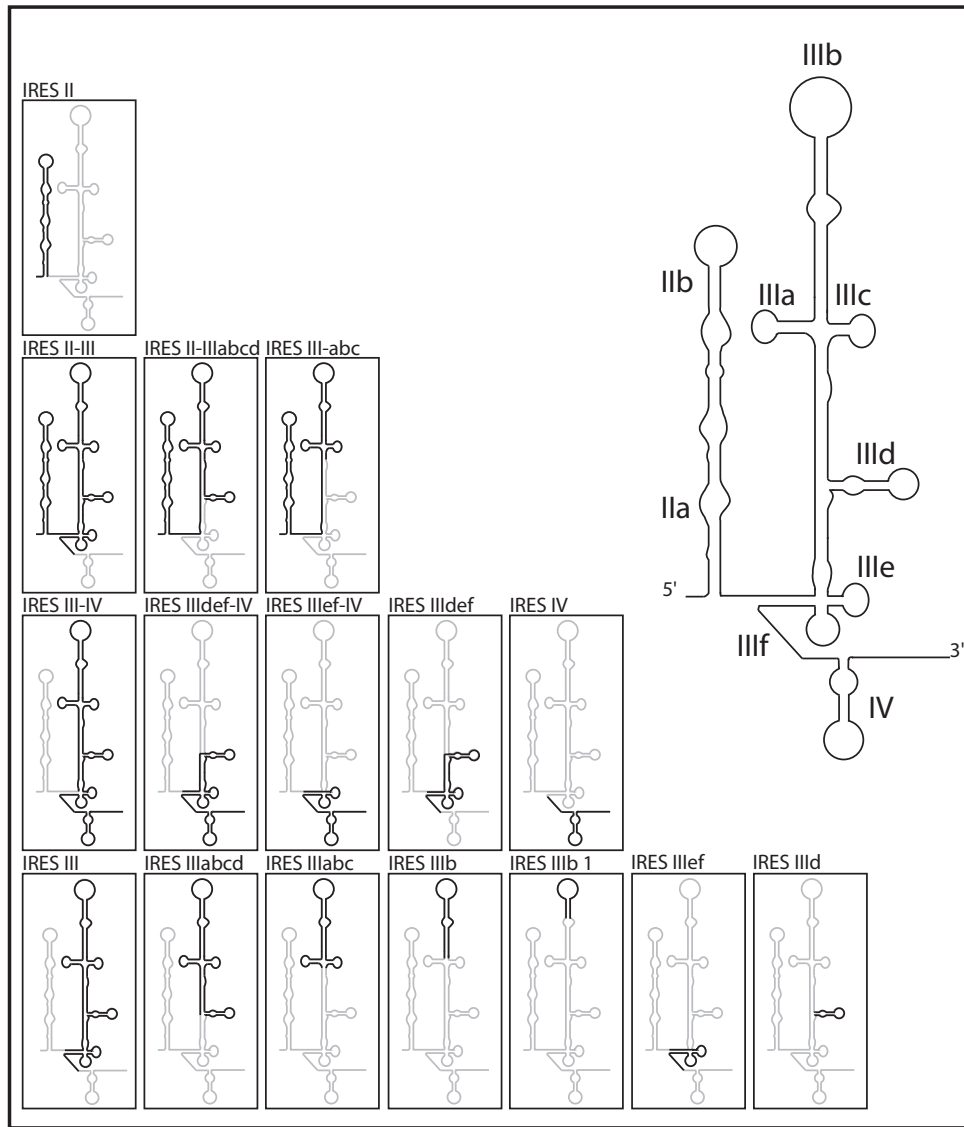


FIG61 : Illustration de la dissection moléculaire réalisée sur l’IRES du VHC.

Toutes les constructions possède 3 guanines en 5' afin d’initier la transcription in vitro et une adénine en 3' correspondant au résidu du site de restriction enzymatique HindIII.

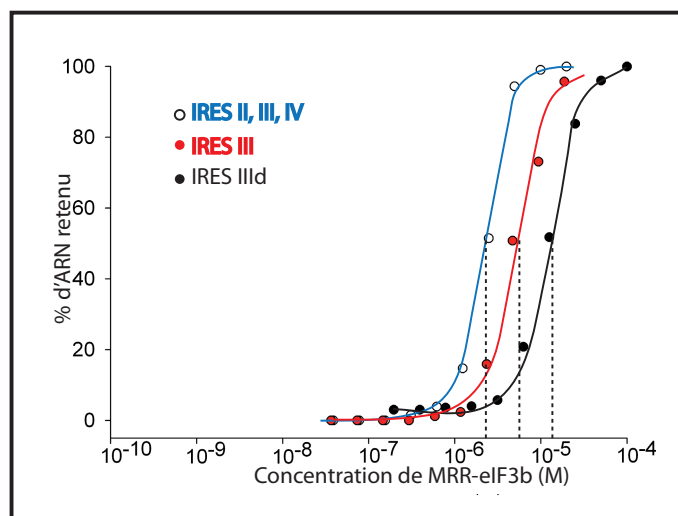


FIG62 : Etude de l’interaction de l’IRES avec le motif MRR-eIF3b.

IRES domaine II-III-IV : K_d = 2μM
 IRES domaine III : K_d = 6μM
 IRES domaine IIIId : K_d = 20μM

tats nous n'avons pas été en mesure d'interpréter plus en détails ce phénomène. Le domaine II est souvent présenté comme impliqué dans la stabilisation de l'IRES.

La **figure 62** présente les résultats d'interaction mesurés entre l'IRES entier, le domaine III ou seulement le petit sous domaine IIIId. D'après ces résultats, nous pouvons donc affirmer que le domaine IIIId est suffisant pour interagir avec le MRR-eIF3b. Cependant, l'affinité mesurée avec le domaine III étant 4 fois supérieure que celle mesurée pour le sous-domaine IIIId, nous pouvons en conclure que l'interaction entre l'IRES et le domaine MRR implique le domaine IIIId mais aussi d'autres nucléotides localisés dans le domaine III. Nous avons utilisé la technique de RMN pour étudier les acides aminés du MRR-eIF3b impliqués dans l'interaction du domaine IIIId. Les résultats sont présentés dans la **figure 63**. Les mêmes acides aminés du MRR-eIF3b sont retrouvés impliqués dans la formation du complexe avec l'ARN précédemment identifiés avec l'IRES entier. Ces résultats confirment que le domaine IIIId représente l'élément de reconnaissance de l'IRES pour le domaine MRR-eIF3b.

Afin d'identifier précisément les nucléotides du domaine IIIId impliqués nous avons utilisé plusieurs techniques comme les empreintes chimiques ou enzymatiques. Cependant après plusieurs tentatives nous ne sommes pas arrivés à identifier clairement des nucléotides. Ces résultats peuvent s'expliquer par le fait que l'affinité entre le motif MRR-eIF3b et l'IRES est relativement faible. Une étude de RMN, actuellement en cours, va permettre d'identifier précisément les nucléotides du domaine IIIId impliqués dans cette interaction.

Ne pouvant attendre ce modèle final et disposant de données de diffusion permettant de modéliser le complexe IIIabcd / MRR-eIF3b, j'ai procédé à la modélisation du complexe.

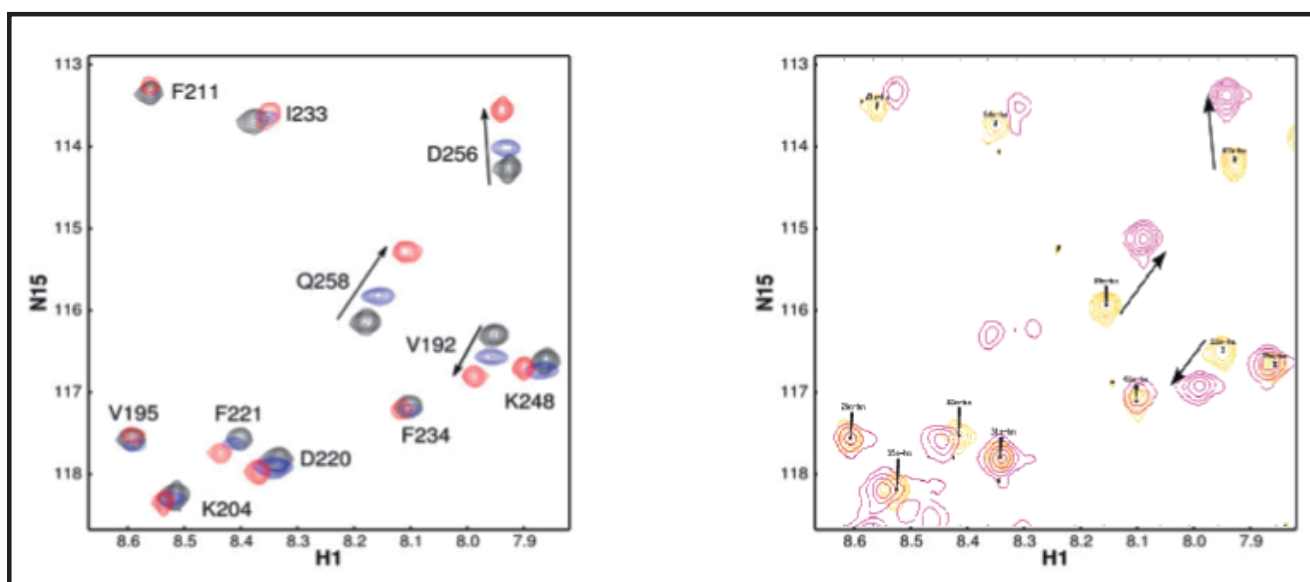


FIG63 : Résultats de l'étude de spectroscopie RMN menée sur l'identification des acides aminés du domaine MRR-eIF3b impliqués dans l'interaction avec l'IRES IIIId.

A : Spectre RMN 2D représentant une région identifiant des déplacements de pics correspondant aux acides aminés impliqués dans l'interaction MRR-eIF3b / IRES II-III-IV.

B : Spectre RMN 2D représentant une région identifiant des déplacements de pics correspondant aux acides aminés impliqués dans l'interaction MRR-eIF3b / IRES IIIId.

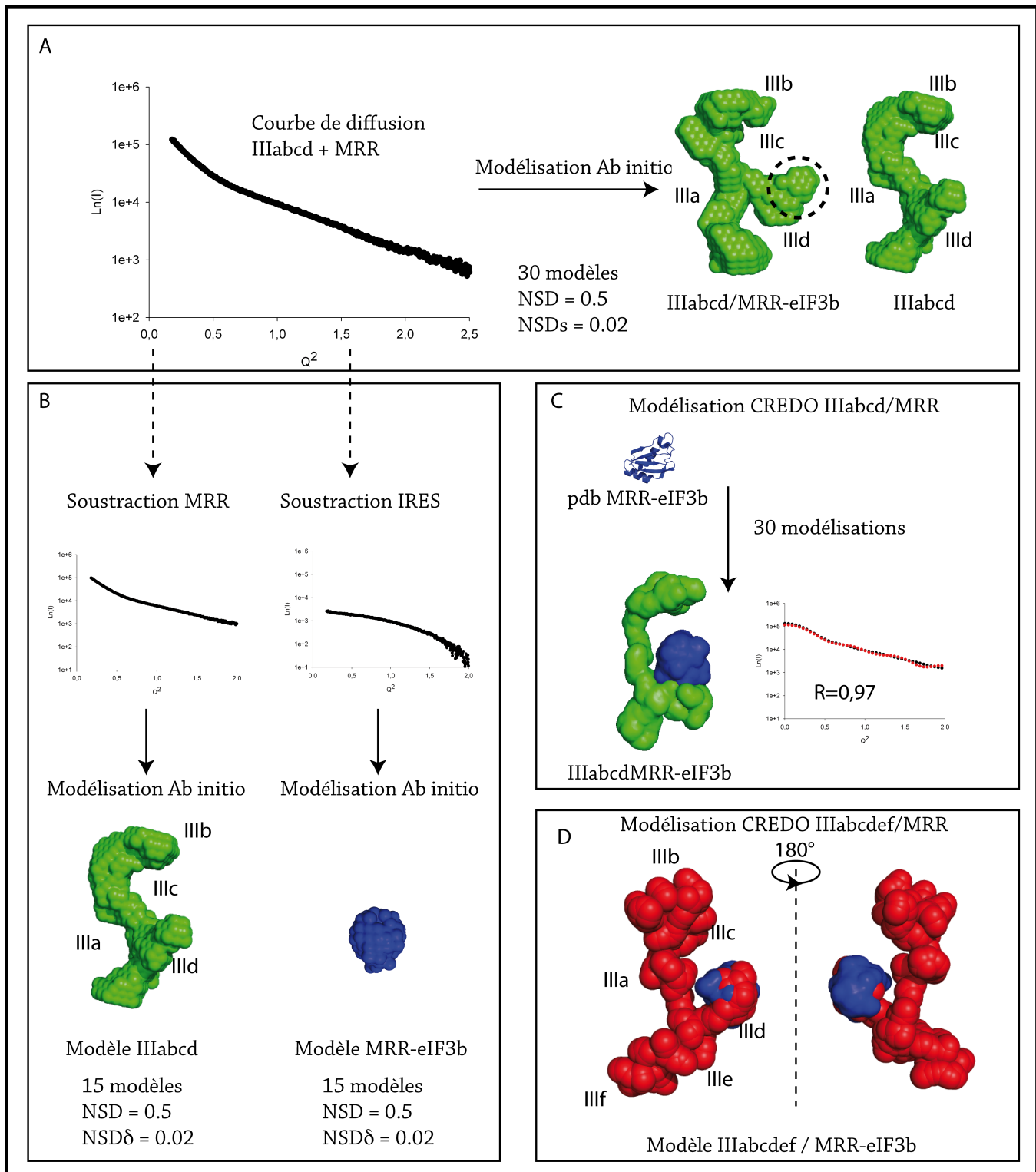


FIG64 : Etude par SAXS de l'interaction entre le domaine IIIId et le motif MRR-eIF3b.

A : Courbe de diffusion et modélisation ab initio du complexe IIIabcd / MRR-eIF3b.

B : Modélisation des deux constituants du complexe IRESIIIabcd/MRR-eIF3b.

C : Modélisation par reconstruction en «rigid body» par le programme CREDO du complexe IRESIIIabcd/MRR-eIF3b.

D : Modélisation par reconstruction en «rigid body» par le programme CREDO du complexe IRESIIIabcdef/MRR-eIF3b.

V Etude structurale du complexe MRR/IRES IIIabcd par SAXS

Afin de caractériser la forme du complexe entre l'IRES du VHC et le motif MRR-eIF3b, nous avons réalisé une étude structurale par SAXS. Les 7 fragments qui ont été précédemment utilisés pour l'étude structurale de l'IRES en solution ont aussi servi à la formation du complexe avec le domaine MRR. Je disposais donc des courbes de diffusion de 7 complexes ainsi que de la courbe de diffusion de la protéine MRR-eIF3b seule. Après analyse des données, seuls les complexes IIIabcd/MRR et IIIabcde/MRR, ont pu être utilisés pour une modélisation. Le domaine MRR étant de petite taille (10kDa), et l'ARN étant de forme étendue et de grande taille, la localisation du domaine MRR n'a pas été possible pour les autres constructions. Dans un premier temps, j'ai réalisé une modélisation *Ab initio* du complexe IIIabcd/MRR-eIF3b. Comme le montre les résultats de la **figure 64A** on constate que l'on retrouve la forme du domaine IIIabcd avec la présence d'une densité supplémentaire localisée proche du domaine IIIcd. Afin de confirmer ce résultat, j'ai utilisé différentes approches. Tout d'abord j'ai soustrait la courbe de diffusion du domaine MRR-eIF3b à la courbe de diffusion du complexe MRR-eIF3b/IRESIIIabcd à l'aide du programme PRIMUS. Le résultat de cette opération correspond à la diffusion du domaine IIIabcd seul. J'ai procédé de la même façon afin d'obtenir la courbe de diffusion de la protéine seule. Après ce traitement, j'ai réalisé une modélisation *Ab initio* de ces deux molécules par le programme DAMMIN (**FIG64B**). On peut remarquer que l'on retrouve la forme de l'IRES IIIabcd ainsi que la forme de la protéine. Ce résultat indique que la protéine et l'ARN ne change globalement pas de forme lors de la formation du complexe. Afin de poursuivre mes investigations j'ai utilisé le programme CREDO qui a été décrit dans le chapitre II des résultats lors du traitement des données de diffusion de la protéine eIF3b. Le programme, en utilisant la courbe de diffusion du complexe MRR-eIF3b/IRESIIIabcd, a pu reconstruire l'enveloppe correspondant à l'ARN à partir de la structure pdb de la protéine (**FIG64C**). La courbe d'affinement indique un coefficient de corrélation de 0.97, ce qui semble compatible avec la qualité des données expérimentales. J'ai utilisé la même approche pour toutes les autres constructions d'ARN disponibles. Cependant la protéine est de petite taille (10kDa) et ne permet pas la reconstruction fiable de l'ARN. Cependant, je suis arrivé à générer une enveloppe à partir de la courbe de diffusion du complexe IIIabcde/MRR-eIF3b (**FIG64D**). Ces résultats indiquent que le domaine MRR se localise au niveau du domaine IIIcd. Néanmoins, la technique de SAXS ne permet pas d'obtenir de modèle à haute résolution du complexe IIIcd/MRR-eIF3b. Pour ce faire, il faudra attendre la modélisation du complexe actuellement en cours de raffinement par l'équipe du laboratoire de LRMN à l'IBS.

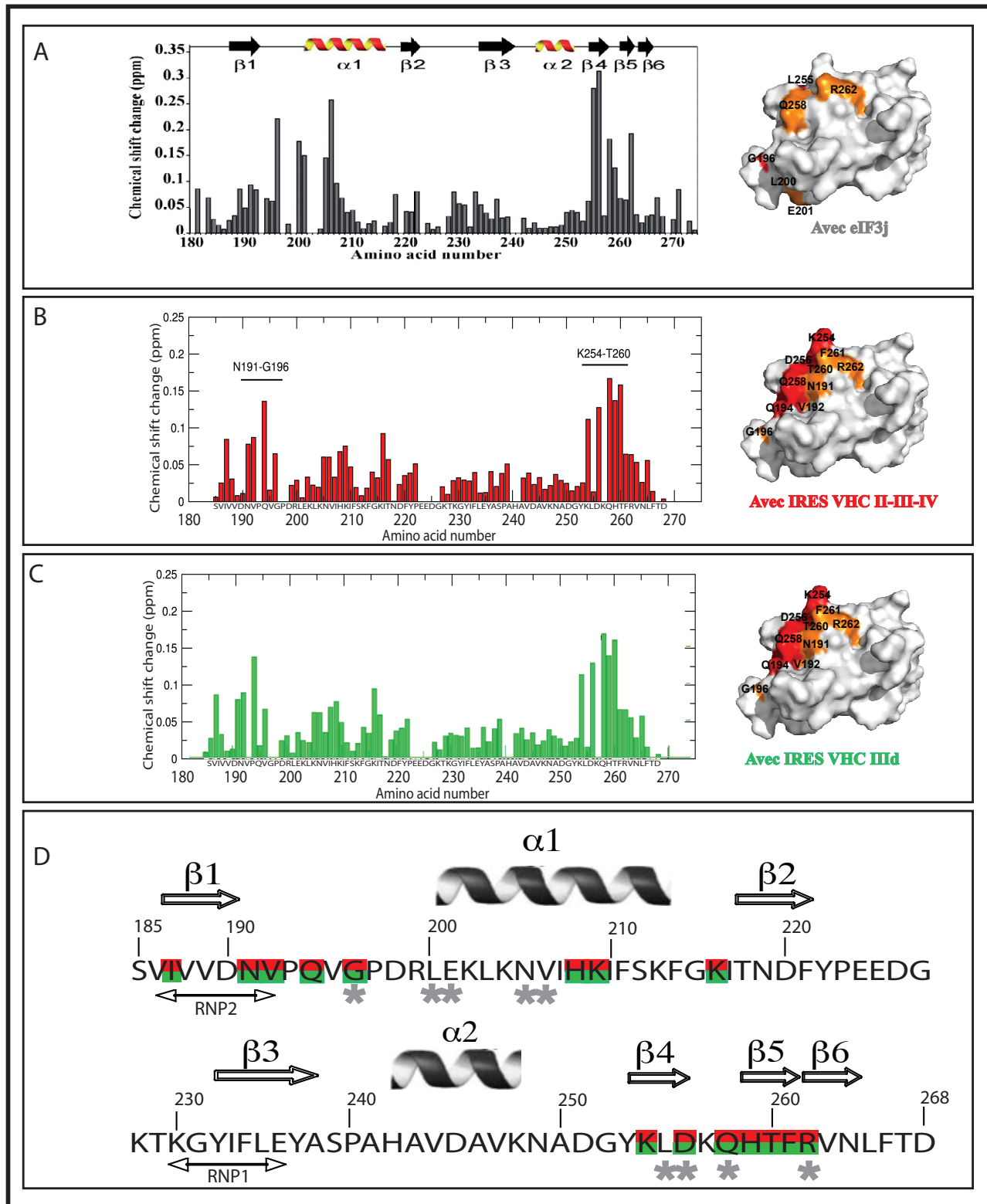


FIG65 : Comparaison des résidus du MRR-eIF3b impliqués dans l'interaction avec la protéine eIF3j ou l'IRES du VHC.

A : Déplacement chimique des acides aminés du domaine MRR-eIF3b en présence de la protéine eIF3J.

B : Déplacement chimique des acides aminés du domaine MRR-eIF3b en présence de l'IRES du VHC II-III-IV.

C : Déplacement chimique des acides aminés du domaine MRR-eIF3b en présence de l'IRES du VHC III d.

D : Représentation des acides aminés impliqués dans l'interaction avec eIF3j (*) ou l'IRES II-III-IV (rouge) ou III d (vert).

VI Conclusion

La combinaison de différentes techniques biochimiques et structurales nous a permis de démontrer que le domaine MRR de la sous-unité b du complexe eIF3 interagissait avec l'IRES du VHC au niveau du domaine IIIId. La technique RMN nous a permis d'identifier un motif situé dans la partie C-terminale, localisé entre les feuillets 4 et 5 et comprenant la boucle 5 qui est impliquée fortement dans l'interaction avec l'ARN viral. Ce motif possède une certaine similitude avec le domaine RRM3 de la protéine PCBP (poly(rC) binding protein). Ce motif contient 5 boucles qui sont impliquées dans l'interaction avec l'ARN (Simpson *et al.* 2004). De plus PCBP aurait la capacité de se lier à l'IRES du VHC (Spangberg et Schwartz 1999).

Une étude publiée en 2007 (ElAntak *et al.* 2007) indique cependant que le motif de reconnaissance à l'ARN, porté par la sous-unité b, n'est pas capable de se lier à l'ARN et qu'il serait exclusivement dédié à des interactions protéine/protéine (MRR non canonique). Les auteurs ont étudié ce domaine par RMN et ont montré que ce motif interagissait avec le domaine N-terminal de la sous-unité j du complexe eIF3.

Dans leur publication, l'étude par RMN indique que les aa : G196, L200, E201, N205, V206, L255, D256, Q258, H259 et R262 de eIF3b sont impliquées dans l'interaction avec eIF3j (**FIG65**). Il est surprenant de constater que ces aa sont situés dans la même région que ceux qui interagissent avec l'ARN. Les auteurs indiquent que l'interaction mesurée entre les deux protéines (MRR-eIF3b et eIF3j) est de 20 μ M, ce qui est 10 fois plus faible que le Kd préalablement déterminé entre l'IRES et le motif MRR. Le MRR posséderait donc une meilleure affinité pour l'IRES qu'envers la protéine eIF3j. Afin de vérifier si cette interaction était exclusive, j'ai produit sous forme recombinante la protéine eIF3j. La protéine a été purifiée après deux étapes de chromatographie, affinité (HisSelect) et exclusion de taille (S75) (**FIG66**). J'ai essayé, sans succès de former le complexe eIF3j/MRR-eIF3b dans mes conditions expérimentales. Le complexe eIF3j/MRR-eIF3b ne peut exister que dans un tampon contenant moins de sel (ElAntak *et al.* 2007).

Puis, j'ai ajouté de petites quantités d'ARN afin de former un complexe tripartite entre les protéines et l'IRES. Après analyse des résultats par la technique EMSA, il s'est avéré que le domaine MRR précipitait en absence de sel. Il semble donc difficile d'étudier ce complexe dans ces conditions. Ayant toujours l'intention d'élucider ce mécanisme et d'étudier l'effet de la présence de la sous-unité j sur la capacité de fixation du domaine MRR à l'IRES, j'ai utilisé une autre ap-

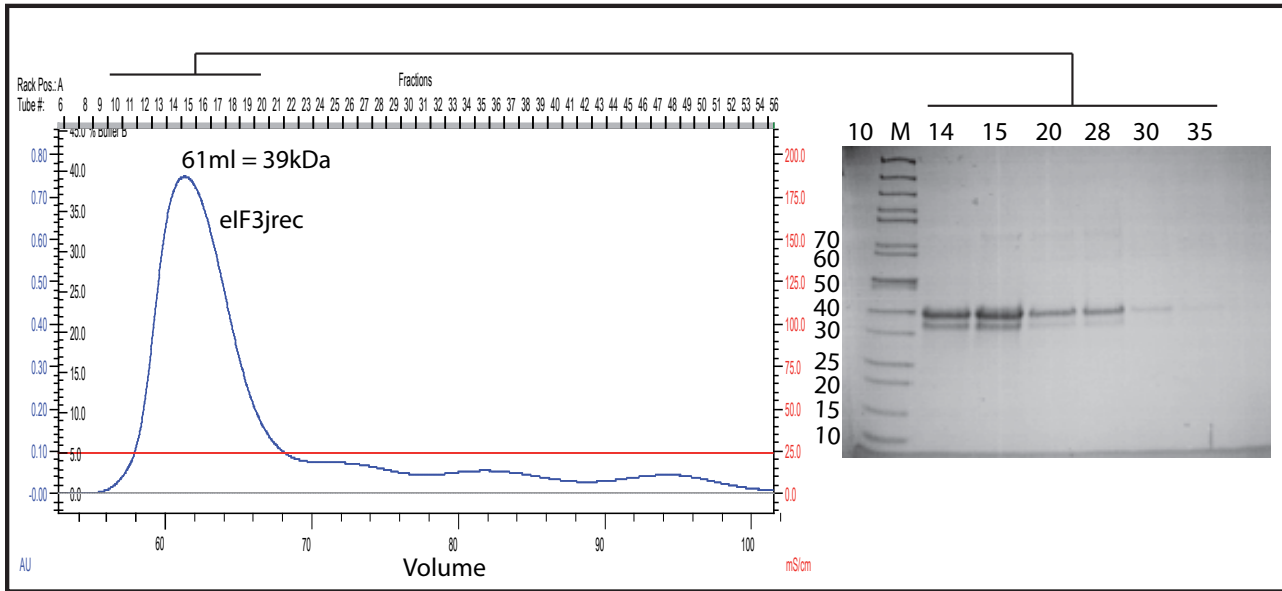


FIG66 : Purification de la protéine eIF3j sous forme recombinante.

Chromatogramme obtenu après purification par filtration sur gel S75 (GEHealthcare) préparative (120ml). Le volume d'exclusion de la protéine est de 61ml, ce qui correspond à une protéine de 39kDa en fonction de la calibration de cette colonne. Gel SDS PAGE à 12% (Laemmli) des fractions de la protéine purifiée. M : Marqueur de poids moléculaire.

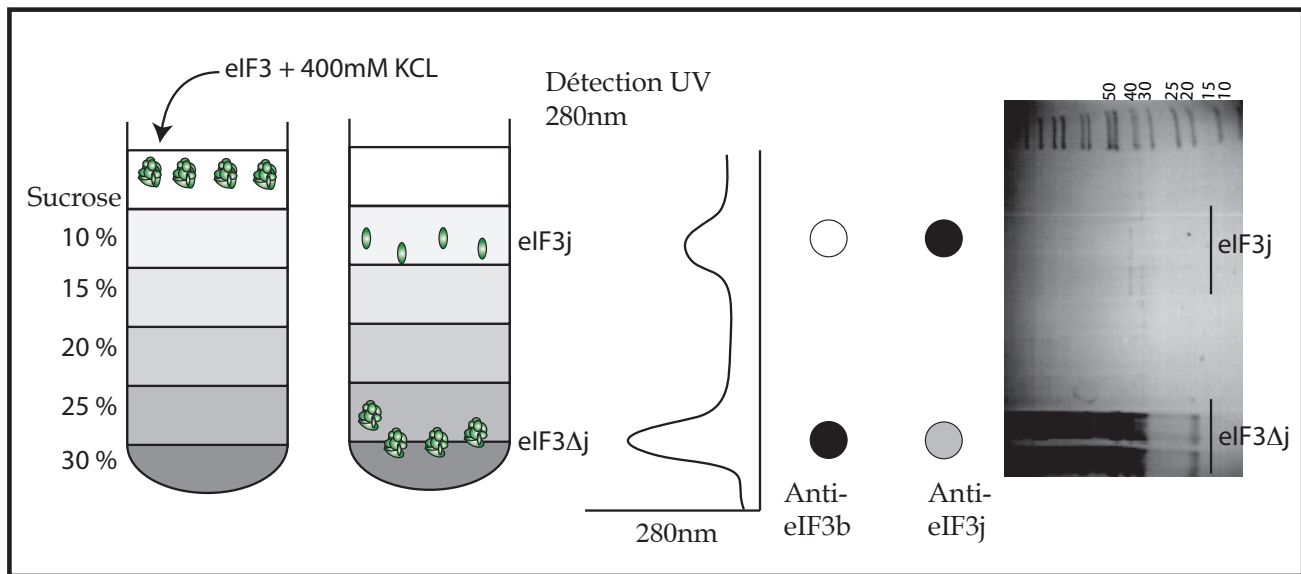


FIG67: Formation du complexe eIF3Δj.

Le complexe eIF3 est déposé sur un gradient de sucrose en présence de 400mM de KCl. Après centrifugation on retrouve la sous unité j dans le haut du gradient et le complexe eIF3Δj dans le bas du gradient. La détection de la présence de protéine est réalisée à 280 nm. Les échantillons sont ensuite déposés sur une membrane de nitrocellulose pour révéler la présence des protéines par l'utilisation d'anticorps spécifiques.

Tout d'abord, il est important de noter que la sous-unité j est souvent retrouvée absente du complexe eIF3 lors de sa purification (Unbehaun *et al.* 2004; Fraser *et al.* 2009).

Les auteurs utilisent la technique de centrifugation sur gradient de sucrose en présence de 400mM de KCl pour éliminer cette sous-unité. J'ai donc utilisé le même procédé pour dissocier la sous-unité j du complexe. Le facteur eIF3 purifié est déposé sur un gradient de sucrose 10-30%. Après centrifugation, la sous-unité j est retrouvée en haut du gradient alors que le complexe eIF3 Δ j est situé dans le bas du gradient (**FIG67**). Afin de m'assurer de la dissociation de la sous-unité j, j'ai réalisé un gel SDS PAGE de chaque fraction du gradient où l'on voit la sous-unité j au centre du gradient. A l'aide d'anticorps (cf. Mat&Met), j'ai pu suivre indépendamment le complexe eIF3 Δ j ainsi que la sous-unité j. Après avoir éliminé la sous-unité endogène du complexe eIF3 nous avons ré-introduit la sous-unité j sous forme recombinante. J'ai, avec succès, reformé le complexe eIF3 Δ j/eIF3jrec en utilisant la même approche que celle utilisée pour éliminer la sous-unité endogène (**FIG68**). J'ai mis en contact la sous-unité j et le complexe eIF3 Δ j dans un tampon contenant peu de sel (50mM KCl). Après centrifugation, l'excès de sous-unité j recombinante est retrouvée dans le milieu du gradient tandis que le complexe nouvellement formé est retrouvé dans le fond du gradient. Des analyses sur gel SDS PAGE et par western blot valident les résultats, et nous indiquent la présence du complexe eIF3 Δ J/eIF3jrec dans le bas du gradient.

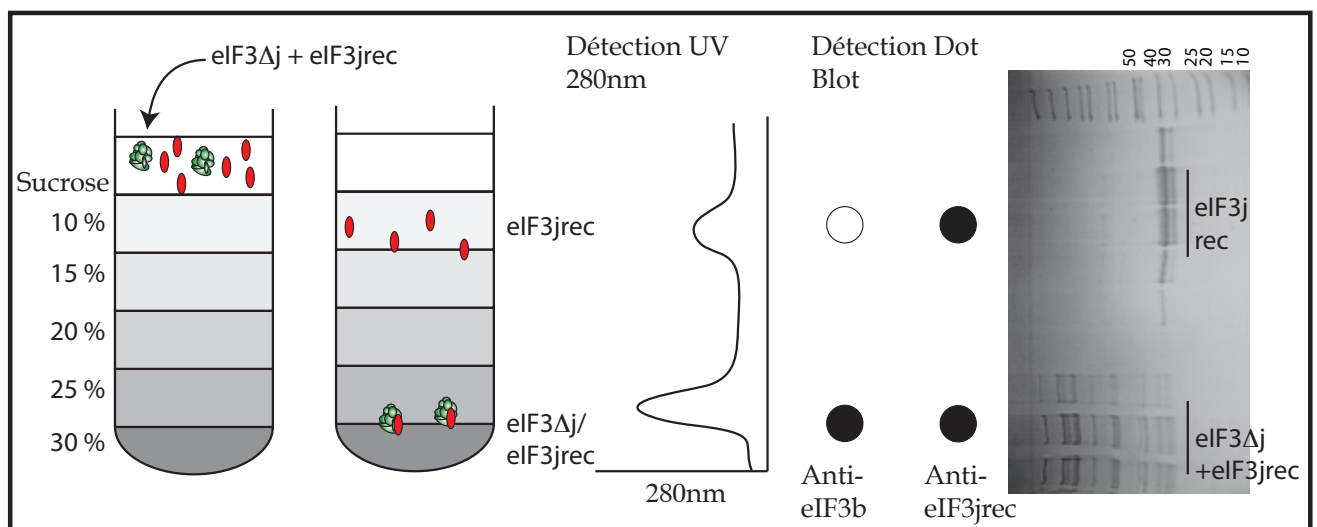


FIG68 : Formation du complexe eIF3 Δ j/eIF3jrec.

Le complexe eIF3 Δ j/eIF3jrec est déposé sur un gradient de sucrose en présence de 50mM de KCl. Après centrifugation on retrouve la sous-unité j rec en excès dans le haut du gradient et le complexe eIF3 Δ j/eIF3jrec dans le bas du gradient.

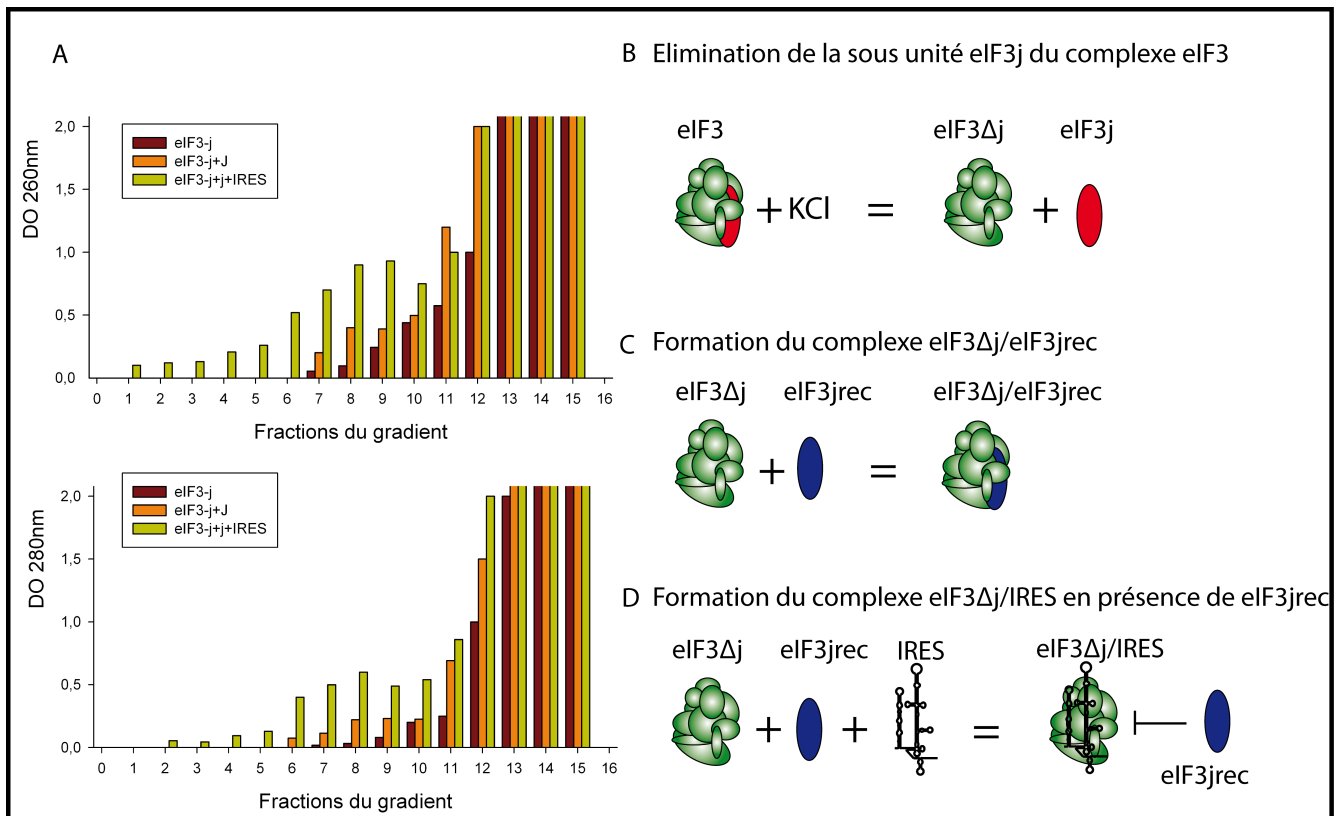


FIG69 : Mise en évidence expérimentale de la formation d'un complexe eIF3Δj/IRES en présence eIF3jrec.

A : Mesure de densité optique à 260 et 280nm sur les différentes fractions du gradient.

B : Elimination de la sous unité eIF3j du complexe eIF3. La sous-unité j est dissociée de eIF3 en présence de 400mM KCl.

C : Formation du complexe eIF3Δj/eIF3jrec. 6μg de complexe eIF3Δj et mise en présence de 20μg de sous-unité eIF3jrec dans un tampon en présence de 50mM de KCl. Le mélange est traité de la même façon que le complexe eIF3. La sous-unité J en excès est retrouvée dans le milieu du gradient, le complexe eIF3Δj/eIF3jrec nouvellement formé et retrouvé dans le bas du gradient.

D : Test de compétition entre l'IRES et la sous-unité eIF3jrec sur le complexe eIF3Δj. eIF3Δj et mise en présence de la sous-unité eIF3j et de l'IRES en présence de 50mM de KCl. Le mélange est traité de la même façon que le complexe eIF3. La sous-unité J est retrouvée dans le milieu du gradient, tandis que l'IRES est retrouvé dans le bas du gradient associé à eIF3.

Après avoir éliminé la sous-unité j du complexe eIF3 et validé la formation du nouveau complexe eIF3 Δ j/ eIF3jrec *in vitro*, j'ai testé la capacité de la sous-unité j à perturber la formation du complexe eIF3 Δ j/IRES. J'ai tout d'abord formé le complexe eIF3 Δ j/IRES *in vitro*, avant d'ajouter dans un rapport molaire, la sous-unité eIF3jrec (**FIG69**). Des mesures d'absorbance à 280 et 260 nm indiquent la présence de l'IRES dans les mêmes fractions que le complexe eIF3 Δ j. L'IRES se lie donc au complexe eIF3 Δ j indiquant que la sous-unité j n'est pas requise pour la liaison de l'IRES sur eIF3. De plus, nos résultats semblent indiquer que la présence de l'IRES sur eIF3 empêche la fixation de la sous-unité eIF3jrec. En effet, on retrouve une grande majorité de la sous unité j recombinante au centre du gradient. D'après les données de la littérature, la sous-unité j interagit avec 40S et stabilise le complexe eIF3/40S dans un système coiffe dépendante et se localise dans le sillon de l'ARNm (Fraser *et al.* 2004; Fraser *et al.* 2009). Le domaine II de l'IRES du VHC se localise aussi dans cette région, ce qui implique que la liaison IRES/ 40S est exclusive de eIF3. Nos résultats indiquent qu'en présence de la sous-unité j, le motif MRR serait lié par cette sous-unité et ne pourrait interagir avec l'IRES. Dans ce cas, la sous-unité j serait nécessaire pour la formation du complexe 40S/eIF3. En revanche, en présence de l'IRES du VHC, le MRR de la protéine eIF3b interagit avec l'IRES et la sous-unité j ne peut plus se fixer sur le MRR-eIF3b. L'absence de eIF3j dans la stabilisation du complexe 40S/eIF3 est remplacée par la présence de l'IRES du VHC (Fraser *et al.* 2008).

CONCLUSION & PERSPECTIVES

Nous avons montré, en utilisant la technique de SAXS, que la partie 5' de l'ARN du VHC était de forme allongée (200Å) en solution. L'approche de modélisation par traitement des données de diffusion de lumière nous a permis d'étudier en détail cette structure et complète les différents modèles publiés à ce jour (Kieft *et al.* 2008). La nature dynamique de l'ARN et la taille de cette molécule a rendu son étude difficile par des techniques conventionnelles. L'approche par SAXS nous a conduit à assembler les différents domaines de l'ARN entre eux et à identifier des domaines encore non étudiés à ce jour, comme le domaine III_f. Nous avons été en mesure d'introduire toutes les structures à haute résolution disponibles à ce jour dans l'enveloppe de l'IRES, permettant de raffiner le modèle actuel et de localiser individuellement chaque sous-domaine l'un par rapport à l'autre. De plus, cette technique nous a permis de modéliser les nucléotides formant la boucle apicale du domaine III difficile à étudier du fait de sa grande flexibilité. Dans un dernier point, nous avons étudié la dynamique du domaine II qui semble pouvoir s'orienter à 180° par rapport au domaine III. Ces données confirment de précédents résultats et indiquent que le domaine II est orienté différemment lorsque l'IRES est fixé à 40S. Afin de poursuivre ces travaux, il serait intéressant de tester si la rigidification des zones flexibles identifiées dans cette étude inhibe la liaison de l'IRES pour 40S.

Afin d'étudier le rôle de la protéine eIF3b ainsi que de son domaine MRR, nous avons produit sous forme recombinante différents domaines protéiques. La cristallographie aux rayons X de ces deux protéines ayant échoué, nous avons utilisé la technique de SAXS qui nous a permis d'étudier la forme et la taille de la protéine eIF3b humaine. Cette étude présente pour la première fois l'enveloppe de cette protéine et nous indique la présence d'une molécule de forme allongée présentant plusieurs bourgeonnements à ses extrémités. Afin d'améliorer ce modèle, j'ai conduit plusieurs études bioinformatiques qui m'ont permis d'identifier une région centrale potentiellement structurée qui a été introduite dans l'enveloppe de SAXS. J'ai aussi pu localiser le domaine MRR dans cette enveloppe qui se retrouve ainsi positionné dans une région accessible au solvant. Ces travaux ont permis d'appréhender pour la première fois la structure de la protéine eIF3b sous forme entière en solution.

L'étude de la formation du complexe IRES/eIF3 nous indique une constante de dissociation de 10nM en accord avec la valeur de 35nM de la littérature (Kieft *et al.* 2001). Lorsque eIF3b est utilisé, la constante de dissociation pour l'ARN tombe à 2µM. Cependant, nous avons montré que cette association est bien spécifique. La différence d'affinité mesurée entre la protéine eIF3b et le complexe eIF3 provient certainement du fait que 4 autres sous-unités du facteur eIF3 (a, c, d et f) sont également impliquées dans la liaison à l'IRES. La technique de RMN nous a permis de montrer que 9 résidus appartenant au domaine MRR en position N-Terminale de eIF3b étaient impliqués dans la liaison à l'ARN. D'autre part, des expériences de FBA alliées à la technique de RMN nous indiquent que l'élément minimum de l'ARN impliqué dans la liaison est le domaine III_d. Ce domaine composé d'une structure tige/boucle a été résolu par RMN (Lukavsky *et al.* 2000). Des travaux antérieurs utilisant des sondes

chimiques avaient montré que le domaine III était le domaine principalement impliqué dans l'interaction avec eIF3 (Sizova et al. 1998). Plus précisément se sont les nucléotides de la partie apicale du domaine III, soit le fragment IIIabc qui semble le plus réactif. Il est possible que ces nucléotides soient engagés dans la liaison avec les 4 autres sous-unités de eIF3. Nous étudions par RMN après marquages isotopique de l'ARN ($^{15}\text{N}^{13}\text{C}$), les nucléotides du domaine IIIId impliqués dans la fixation à eIF3b-MRR. Les résultats ne sont que préliminaires mais semblent cependant impliquer quelques nucléotides de la tige du domaine IIIId.

Il est remarquable de noter que ce domaine IIIId représente le site de fixation de 40S et semble indispensable à l'initiation de la traduction de manière indépendante de la coiffe (Lukavsky *et al.* 2009) (**FIG70**). Nous pouvons remarquer que ce fragment est très largement conservé chez les Pestivirus et les Falivirus et nécessaire au fonctionnement de l'activité IRES (Kieft *et al.* 2000). Les auteurs ont identifié les 3 guanines situées dans la partie centrale de la boucle du domaine IIIId comme impliqué dans le processus de recrutement de la sous-unité 40S (Lukavsky *et al.* 2009). En effet les mutations introduites dans ce motif réduisent significativement l'initiation conduite par la séquence IRES. La structure de ce domaine ayant été résolue par RMN (Lukavsky *et al.* 2000), les auteurs ont pu identifier un motif E-loop qui serait impliqué dans l'interaction avec la sous-unité 40S. De récents travaux (Babaylova *et al.* 209) confirment que les 3 guanines du domaine IIIId de l'IRES sont impliquées dans la liaison avec la sous-unité 40S

Le domaine IIIId pourrait représenter une plateforme de recrutement de la sous-unité 40S et du facteur eIF3. Ce domaine pourrait, à l'image de la sous-unité eIF3j, faire le lien entre ces deux acteurs de l'initiation et participer à la stabilisation du complexe IRES/40S/eIF3. Cette hypothèse a également été émise par l'équipe du Pr. Doudna (Fraser *et al.* 2009).

Toutefois, le domaine MRR-eIF3b a été préalablement identifié comme impliqué dans la formation d'un complexe avec la sous-unité eIF3j dans un contexte coiffe dépendante. Ces résultats suggèrent la présence d'un mécanisme complexe qui pourrait impliquer la sous-unité eIF3j comme senseur d'une activité IRES. En effet, la sous-unité eIF3j pourrait faire la jonction entre 40S et eIF3 (interaction avec le MRR de eIF3b) dans un contexte coiffe dépendante. En revanche, dans un contexte coiffe indépendante, le domaine MRR-eIF3b interagit avec l'IRES il ne peut donc pas s'associer à eIF3j. Ainsi, la sous-unité j serait la protéine responsable du choix de la voie coiffe dépendante ou coiffe indépendante. Evidement de telles hypothèses devront être vérifiées ultérieurement.

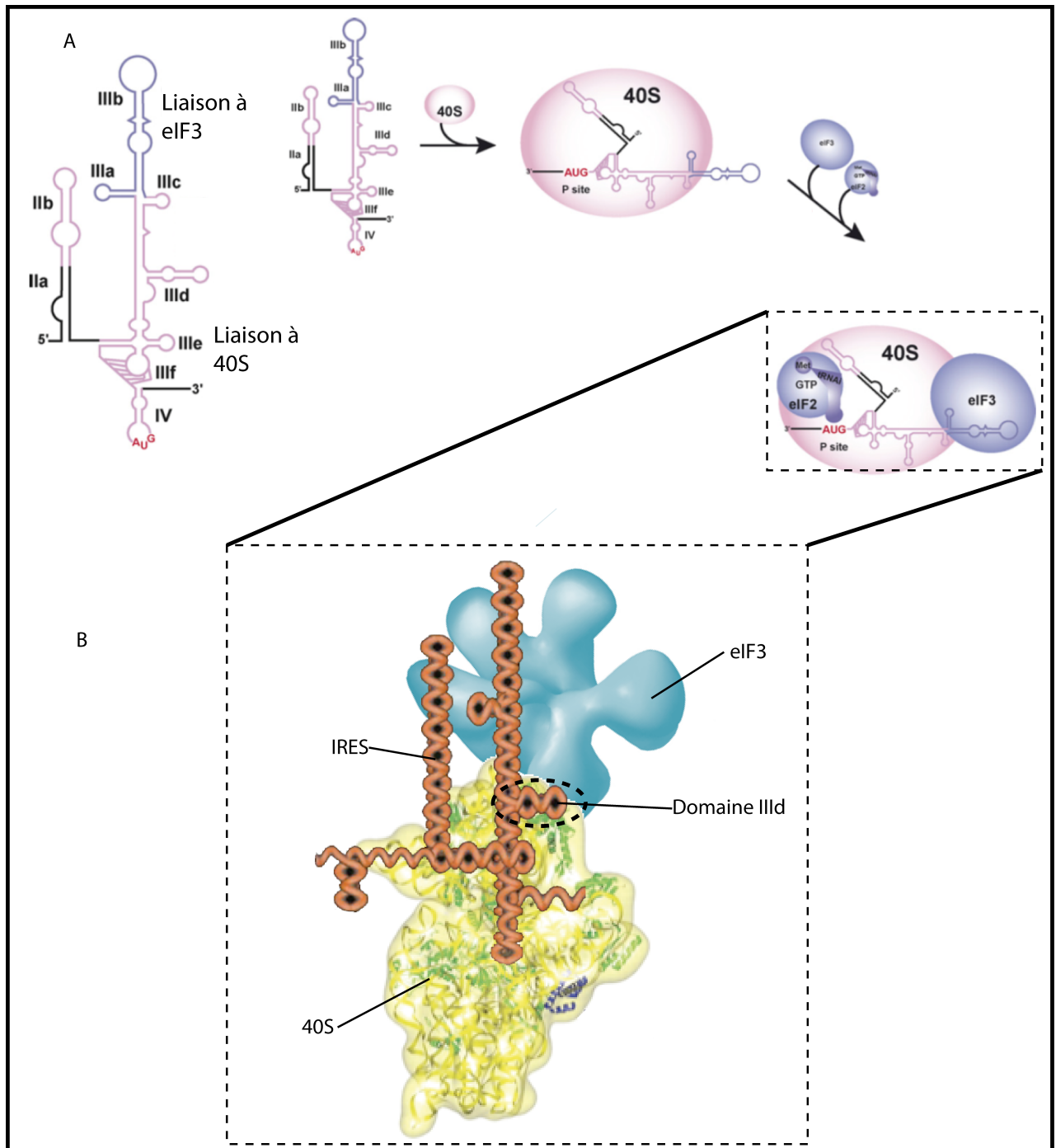


FIG70 : Recrutement de la sous-unité 40S et du facteur eIF3 par l'IRES du VHC.

A : Représentation schématique du recrutement des différents acteurs de l'initiation coiffe indépendante conduite par l'IRES. En rose, les domaines contenant des nucléotides impliqués dans une interaction avec la sous-unité 40S. En violet, les domaines contenant des nucléotides impliqués dans une interaction avec le facteur eIF3, source (Lukavsky et al. 2009).

B : Représentation schématique de la fixation de la sous-unité 40S et du facteur eIF3 sur l'IRES du VHC avec le domaine IIIId comme motif commun d'interaction.

EPILOGUE

Il y a 10 ans, j'ai eu la chance d'effectuer un stage de BTS dans un laboratoire de biochimie des protéines au CEA-Grenoble où l'ATPase CadA de *Listeria monocytogenes* était le cheval de bataille. Le Dr Patrice Catty, que je ne connaissais pas, m'a offert la possibilité de travailler à ses côtés. Ce premier «pas» dans le monde de la recherche, et je ne le savais pas encore, m'avait totalement contaminé. Suite à cette première inoculation, j'ai poursuivi mes études pour mettre à profit mes connaissances techniques dans le monde de l'industrie pharmaceutique. Cependant, les démangeaisons furent de plus en plus fortes, et je ne pus résister à une piqûre de rappel. Je suis donc revenu pour travailler dès que cela fut possible : des vacances, des stages non rémunérés pendant les vacances... bref, j'étais devenu dépendant de la recherche, un véritable accro de la pipette.

Ne pouvant plus me passer de cette drogue, j'ai repris mes études pour atteindre l'ultime dépendance et l'accepter : « oui je suis drogué à la recherche, et alors? »

C'est là, après une petite escapade au Canada pour travailler quelques mois à la Québécoise, que je suis revenu en France pour suivre un cursus scientifique digne de ce nom. Etant loin des laboratoires et des pipettes pendant cette formation, j'ai tout fait pour obtenir une certaine dose annuelle de recherche. N'étant pas chose facile, j'ai souvent du être insistant mais cela me permit de perfectionner de nombreuses techniques en biochimie et en biologie moléculaire. Chemin faisant, j'ai croisé la route du Dr. Jean-Philippe Kleman qui travaillait sur des protéines «passagères» à l'Institut de Biologie Structurale. Lors de cette rencontre, une véritable alchimie finit de confirmer ma dépendance à la recherche et me conduisit jusqu'en Master II. Malheureusement, par manque de chance et de financement, cette interaction ne put perdurer en thèse et je me mis à la recherche d'un autre laboratoire. Jusqu'à ce jour, un vendredi de septembre 2006, je ne sois contacté par téléphone : Cet appel, qui me proposait de travailler à l'UVHCI, émanait du directeur de l'école doctorale Chimie et Science du Vivant via le Pr. Michel Robert Nicoud, le Pr. Rob Ruigrok et du Pr. Emmanuel Drouet. Apparemment, mon cas avait été étudié (suite à quelques manifestations personnelles peut-être ?) et l'on m'offrait une chance d'être contaminé pour de bon (d'avoir « la » dose de recherche ultime : la thèse), une chance inespérée mais tellement souhaitée de rester à la paillasse encore quelques années et de poursuivre mes ambitions. Après un court entretien le lundi, le Dr. Florence Baudin et le Pr. Emmanuel Drouet ont fait le pari, sans me connaître, de me donner un sujet de

thèse ambitieux, complexe et totalement nouveau à mes yeux. C'est à ce moment précis, disposant d'une bourse de thèse de trois ans que ma vie a sombré dans la recherche... N'étant encore qu'un novice, j'ai rassemblé toutes mes connaissances, ma volonté et mes motivations pour vivre ces trois années de recherche au maximum. Ne connaissant rien à la biochimie des ARN j'ai été guidé par Florence. Grâce à ses conseils, j'ai rapidement apprivoisé l'ARN du virus de l'hépatite C, je l'ai disséqué, analysé pour en comprendre sa structure et sa dynamique. J'ai produit des dizaines de protéines pour mieux les appréhender et j'ai combiné ces deux molécules pour en étudier leurs mécanismes. Après ces recherches, ses gloires et ses déboires, ces moments de stress et de satisfaction, je me rends compte du chemin parcouru, d'un technicien, je suis devenu un étudiant chercheur. Ce parcours, en dehors des sentiers battus, m'a permis de construire mes recherches pas à pas. Disposant de bases techniques supplémentaires suite à ma formation initiale, j'ai ainsi pu explorer en profondeur ce sujet passionnant. Ce chemin, si riche en rencontres, en expériences, avec ses doutes et ses changements m'a permis, chaque jour, d'évoluer un peu plus et de considérer mon sujet de thèse sous un angle différent. Il reste bien sûr plus d'une vie de chercheur sur ce sujet, et les étudiants qui suivront devront compléter les résultats qui ont été présentés dans ce manuscrit. Cependant, je suis fier d'avoir participé à ces recherches en espérant avoir contribué, ne serait-ce qu'un peu, à la compréhension de ce mécanisme ingénieux de piraterie de la machinerie cellulaire. J'espère pouvoir continuer mon parcours qui est arrivé à un premier pallier, qui va ainsi me permettre de savourer tout le potentiel de la recherche et de me conduire vers d'autres sujets encore plus passionnants les uns que les autres. Plus que tout, je tenais une dernière fois et de façon solennelle à remercier ces chercheurs et ces professeurs qui m'ont fait confiance et qui m'ont laissé vivre mes recherches au gré des découvertes. Merci à ma famille et à mes amis qui m'ont soutenu tout au long de ce parcours qui touche à sa fin.

A vous tous : Merci

PUBLICATION

Titre : Human initiation factor eIF3 subunit b interacts with HCV IRES RNA through its N-terminal recognition motif.

Une partie des résultats présentée dans ce chapitre ont fait l'objet d'une publication scientifique dans le journal FEBS en 2009. Cette publication avait pour but de mettre en évidence l'implication du motif MRR-eIF3b dans la formation d'un complexe entre l'IRES du VHC et la protéine eIF3b.



Human initiation factor eIF3 subunit b interacts with HCV IRES RNA through its N-terminal RNA recognition motif

Julien Pérard^a, Rodolfo Rasia^b, Jan Medenbach^c, Isabel Ayala^b, Jérôme Boisbouvier^b, Emmanuel Drouet^a, Florence Baudin^{a,c,*}

^a Unit of Virus Host-Cell Interactions, UJF-EMBL-CNRS (UMR 5233), 6 rue Jules Horowitz, 38042 Grenoble Cedex 9, France

^b Institut de Biologie Structurale (IBS) J.P. Ebel, UJF-CEA-CNRS (UMR 5075), 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

^c European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, 69117 Heidelberg, Germany

ARTICLE INFO

Article history:

Received 22 October 2008

Revised 13 November 2008

Accepted 18 November 2008

Available online 6 December 2008

Edited by Gianni Cesareni

Keywords:

eIF3b

HCV IRES RNA

NMR spectroscopy

RNA–protein interaction

Translation initiation

ABSTRACT

Many viral mRNAs contain a 5'-UTR RNA element called internal ribosome-entry site (IRES), which bypasses the requirement of some canonical initiation factors allowing cap-independent translation. The IRES of hepatitis-C virus drives translation by directly recruiting 40S ribosomal subunits and binds to eIF3 which plays a critical role in both cap-dependent and cap-independent translation. However, the molecular basis for eIF3 activity in either case remains enigmatic. Here we report that subunit b of the eIF3 complex directly binds to HCV IRES domain III via its N-terminal-RRM. Because eIF3b was previously shown to be involved in eIF3j binding, biological implications are discussed. © 2008 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

In eukaryotic organisms, translation of most messenger RNAs requires numerous eukaryotic initiation factors (eIFs) that stimulate and regulate formation of an elongation competent ribosome at the translation initiation codon. This includes recognition of the m⁷G-structure, recruitment of the 40S ribosomal subunit charged with Met-tRNA^{Met}, scanning of the 5' untranslated region (UTR) and start codon recognition finally resulting in subunit joining and translation of the mRNA. Initiation of HCV protein synthesis does not require the recognition of an m⁷G-structure. Instead, the HCV RNA genome contains a conserved and well-defined structure present in its 5'-UTR which acts as an internal ribosome-entry site (IRES) [1]. Briefly, the small ribosomal subunit (40S) and eukaryotic initiation factor eIF3 bind specifically to the HCV IRES, allowing direct recognition of the start codon present in the 5'-UTR of the viral mRNAs. Ribosomal 40S subunits bind to the HCV IRES with a very high affinity ($K_D = 1.9$ nM) in the vicinity of the initiation codon [2]. So far, eight of 40S ribosomal subunit proteins have been suggested to interact with the HCV IRES [3,4]. eIF3 is a multiprotein complex (ca 800 kDa) containing 13 subunits in

mammals (named eIF3a to eIF3m; nomenclature reviewed in [5]) controlling the assembly of 40S ribosomal subunits on mRNAs bearing either an IRES or a 5'cap structure. eIF3 was shown to prevent premature association of the 40S and 60S ribosomal subunits, it interacts with other initiation factors involved in start codon recognition, and helps assemble active ribosomes. Despite its importance in both cellular and viral protein synthesis, the structural basis for eIF3 activities and interactions with the translational machinery are only poorly understood. The recent cryo-electron microscopy reconstruction of human eIF3 at ~30 Å resolution revealed a five-lobed architecture [6]. eIF3 was shown to specifically associate with the apical half of domain III of the HCV IRES [7,8]. Six of its subunits (eIF3a, b, c, d, f and g) have been proposed to exhibit RNA-binding activity but none of them has been characterized individually. Among the eIF3 subunits, eIF3b is considered to be the major scaffolding subunit, interacting with eIF3a, g, i, and j [9,10]. The human eIF3b subunit is an 814 amino acid residue protein, whose primary sequence reveals only one identifiable domain, an RNA recognition motif (RRM), located in its N-terminal moiety. In the present paper, using NMR spectroscopy we show that subunit eIF3b is able to bind directly to domain III of the IRES of HCV RNA through its N-terminal RRM. Because eIF3b was previously shown to be involved in eIF3j binding [11], biological implications are discussed.

* Corresponding author. Fax: +49 6221 387 8518.

E-mail address: baudin@embl.de (F. Baudin).

2. Materials and methods

2.1. Plasmid constructs and protein purification

The DNA fragments encoding full-length eIF3b, the RRM of eIF3b (aa 185–268) and the RRM of eIF3g (aa 227–321) were PCR-amplified from cDNA libraries (ImaGenes) and subcloned into a pET28a-TEV vector for the expression of His-tagged fusion proteins. Proteins were expressed in *E. coli* BL21 (DE3) and purified using a His-Select Cobalt column (SIGMA). Double labeled ^{13}C , ^{15}N -RRM was produced in M9 minimum media and purified according to standard protocols. Eluted fractions were pooled and concentrated prior to size exclusion chromatography, then analyzed by SDS-PAGE, see [Supplementary Data](#).

The eIF3 complex was purified from rabbit reticulocyte lysate (RRL, Green Hectares) essentially as described in [12], see [Supplementary Data](#).

Anti eIF3b-RRM and anti eIF3g-RRM monoclonal antibodies were produced using standard techniques (rProteinA sepharose, GE Healthcare).

2.2. RNA synthesis by *in vitro* transcription

RNAs comprising the IRES elements of HCV RNA were transcribed *in vitro* from linearized plasmids using T7 RNA polymerase. Briefly, HCV IRES element containing plasmid templates were linearized with *HindIII* and purified by phenol/ether extraction followed by ethanol precipitation. Transcription was performed in standard buffer using [α - ^{32}P]-UTP and labeled RNAs were then purified by PAGE.

2.3. Binding reactions and filter-binding-assay

Radioactive RNA was first renatured by heating in water at 95 °C for 1 min, cooled down to 25 °C and incubated in buffer A (20 mM Tris-HCl pH 7.5, 200 mM KCl, 2.5 mM MgCl₂, 1 mM DTT). Proteins

were serially diluted (from 100 μM to 0.1 nM) in a final volume of 180 μl , followed by addition of 20 μl of renatured ^{32}P -labeled RNA. After incubation at 37 °C, the reactions were filtered through a nitrocellulose filter and the radioactivity counted. For each series, a blank reaction and total RNA input were determined. Because the amount of labeled RNA is negligible, an apparent dissociation coefficient (K_D) can be estimated from the concentration value at 50% RNA retention. The interference of specific antibodies with complex formation was tested using monoclonal antibodies directed against the eIF3b- or eIF3g-RRMs or against the His-tag. HCV IRES RNA was first prebound to eIF3b-RRM, then incubated with MABs for 30 min at 37 °C using a stoichiometric ratio (1:1). For competition experiments, ^{32}P -labeled HCV IRES RNA was first bound to 20 μM proteins, then increasing concentrations of unlabeled competitor RNA (4.10^{-11} – 4.10^{-6} M) were used to challenge the complex. Reported values are the average of at least three independent experiments.

2.4. NMR studies

The interaction of RRM with IRES was followed by titration of a 1 mM protein solution with a 0.3 mM solution of IRES RNA, both in buffer A at 25 °C. Chemical shift perturbations were followed by acquiring TROSY ^1H - ^{15}N correlation spectra with an 800 MHz spectrometer equipped with a cryogenic probe. Spectra were recorded for the free protein and for protein/ RNA (ratios of 4/1 and 1/1.25). Chemical shift perturbations were calculated as a weighed distance on the ^1H - ^{15}N plane between the signals for the free protein and those of the protein in the presence of excess IRES-RNA, see [Supplementary Data](#).

3. Results and discussion

3.1. eIF3 multisubunit complex binds to the IRES element of HCV RNA

To measure the affinity of eIF3 complex to the IRES present in the HCV RNA, radiolabeled IRES RNA was incubated with increasing

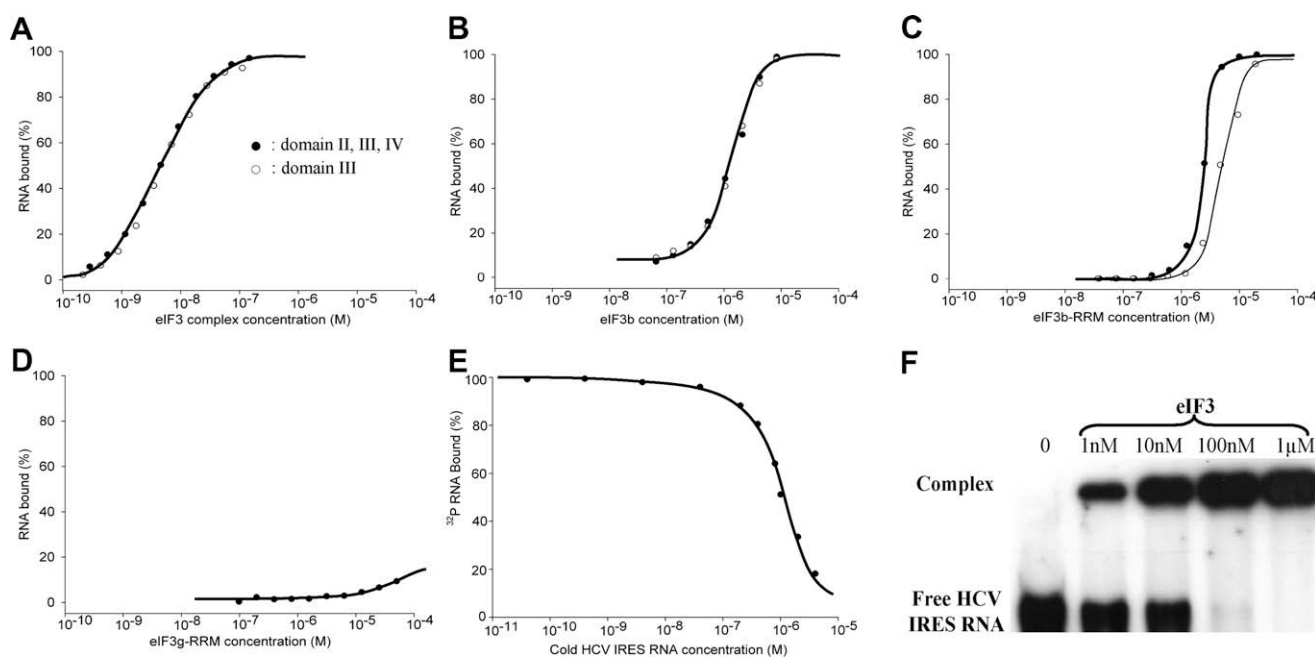


Fig. 1. eIF3 complex, eIF3b and eIF3b-RRM bind to HCV IRES RNA. (A–D) Binding kinetics of rabbit eIF3 complex, human eIF3b, human eIF3b-RRM and human eIF3g-RRM to full-length HCV IRES RNA (●) or domain III only (○) using FBA. (A) eIF3 complex. (B) eIF3b recombinant protein. (C) eIF3b-RRM recombinant domain. (D) eIF3g-RRM recombinant domain. (E) Competitive binding assays. Labeled HCV IRES RNA was first complexed to 2.10^{-5} M of eIF3b-RRM, before addition of increasing concentrations (4.10^{-11} – 4.10^{-6} M) of cold HCV IRES RNA competitor. Results are expressed as percentage of RNA bound. (F) eIF3 binds to HCV IRES RNA in EMSA. Labeled RNA comprising domains II, III and IV was incubated with eIF3 (1 nM–1 μM) and loaded onto a native 6% polyacrylamide gel.

concentrations of purified eIF3 (from 10^{-10} M to 10^{-6} M) and the resulting complexes were quantified by filter-binding-assays (FBA, Fig. 1). We could measure an apparent equilibrium dissociation constant (K_D) of 5 nM for the binding of eIF3 complex to the IRES (Fig. 1A). These results were confirmed by an electrophoretic mobility shift assay (EMSA, Fig. 1F). At 1 nM eIF3 complex concentration, a certain proportion of HCV IRES is already retained in the well due to the high molecular weight of the protein complex. The affinity of eIF3 complex for IRES HCV RNA was already reported in the literature to be ~ 35 nM [2]. The discrepancy between both results (5 nM/35 nM) is certainly not significant since the estimation of the concentration of the eIF3 is achieved on the whole complex and is probably subjected to variations.

3.2. Isolated subunit b of the eIF3 complex binds to the IRES of HCV RNA

In order to test whether subunit eIF3b binds to the IRES element, we expressed and purified human eIF3b. 1D NMR and CD spectra (Supplemental Fig. S4) were recorded to ensure proper folding. Using the same experiments as described above (Fig. 1B), we determined a K_D of 10^{-6} M between eIF3b and HCV IRES RNA (more than 2.5 log above the K_D obtained for the eIF3 multicomplex). This drop of affinity is predictable considering that eIF3b is not the only subunit involved in the RNA binding. Although sub-

units a, d and f of eIF3 were clearly involved in RNA interaction [7], involvement of subunit b was not clearly established [8,13].

3.3. N-terminal recognition motif of eIF3b binds to HCV IRES RNA

Since the N-terminus of eIF3b contains a RNA recognition motif (RRM), between aa 185 and 268 – a sequence often mediating protein–RNA interactions – we tested if this motif was sufficient for binding to the IRES element. As a control, the RRM of eIF3g (aa 239–317) was used. FBA revealed a K_D of approx. 10^{-6} M for the eIF3b-RRM/HCV IRES complex similar to the one measured for the full-length protein (compare Fig. 1B and C). In contrast, the RRM motif derived from eIF3g did not significantly bind to the RNA (Fig. 1D). This result suggests that the RRM motif present in eIF3b is the responsible RNA-binding element. To investigate the stability of the IRES/eIF3b-RRM complex, we performed competition experiments (Fig. 1E). Labeled IRES RNA was first bound to a constant amount of eIF3b-RRM ($2 \cdot 10^{-5}$ M) resulting in complete incorporation of the RNA into the eIF3b-RRM/IRES RNA complex. This complex was then challenged with increasing concentrations of unlabeled HCV IRES RNA (ranging from $4 \cdot 10^{-11}$ M to $4 \cdot 10^{-6}$ M, Fig. 3) confirming the previously determined K_D (compare Fig. 1C and E). Taken together these results show that the RRM motif of eIF3b is sufficient for mediating binding to the IRES element of HCV RNA.

To further characterize the specificity of the binding of the IRES to the RRM of eIF3b, we performed FBA experiments in the presence of increasing concentrations (from $3 \cdot 10^{-7}$ M to $3 \cdot 10^{-5}$ M) of monoclonal antibodies (MAbs) raised against the RRM of either eIF3b or eIF3g or raised against the His-Tag (Fig. 2). Whereas the α eIF3g-RRM and α 6xHis did not affect RNA binding to eIF3b-RRM, the monoclonal α eIF3b-RRM antibody efficiently prevented complex formation.

3.4. The RRM of eIF3b binds to domain III of the HCV IRES element

The secondary structure of HCV IRES RNA consists of four domains named I–IV (Fig. 3A). The affinity of eIF3, eIF3b and eIF3b-RRM for these different domains was determined using FBA as described above (Figs. 1A, B and 3B). eIF3 was able to bind to domain III with an apparent K_D of 5 nM, identical to full-length IRES RNA (Fig. 1A). eIF3b and eIF3b-RRM were also able to associate with domain III as tightly as with the full-length IRES RNA ($K_D = 2 \mu$ M), demonstrating that domain III is necessary and sufficient for eIF3b-RRM binding. eIF3b-RRM was not able to bind either domain II or IV (Fig. 3B).

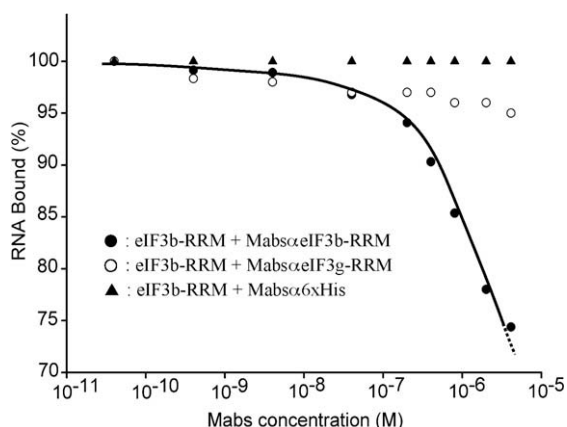


Fig. 2. Specific monoclonal antibodies disrupt the HCV/eIF3b-RRM complex. $2 \cdot 10^{-5}$ M of eIF3b-RRM was incubated with increasing concentrations of specific MAbs α eIF3b-RRM (or α eIF3g-RRM (●) or α 6xHis (▲) prior addition of 32 P-labeled HCV full-length RNA. Results are expressed as percentage of RNA bound.

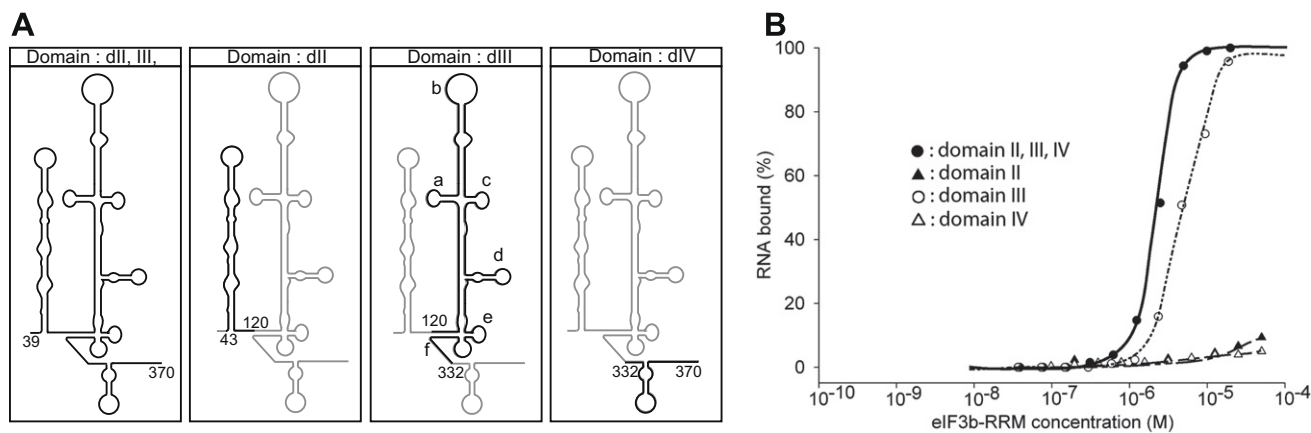


Fig. 3. eIF3b-RRM binds to domain III of HCV RNA. (A) Molecular dissection of HCV IRES-RNA: full-length domains II, III, IV (nt 39–370 of genotype 1b), domain II (nt 43–120), domain III (subdomains abcdef, nt 120–332), domain IV (nt 332–370). (B) Binding kinetics of purified protein eIF3b-RRM to HCV IRES-RNA domains by filter-binding-assays. (●): domains II, III, and IV, (▲): domain II, (○): domain III, (△): domain IV.

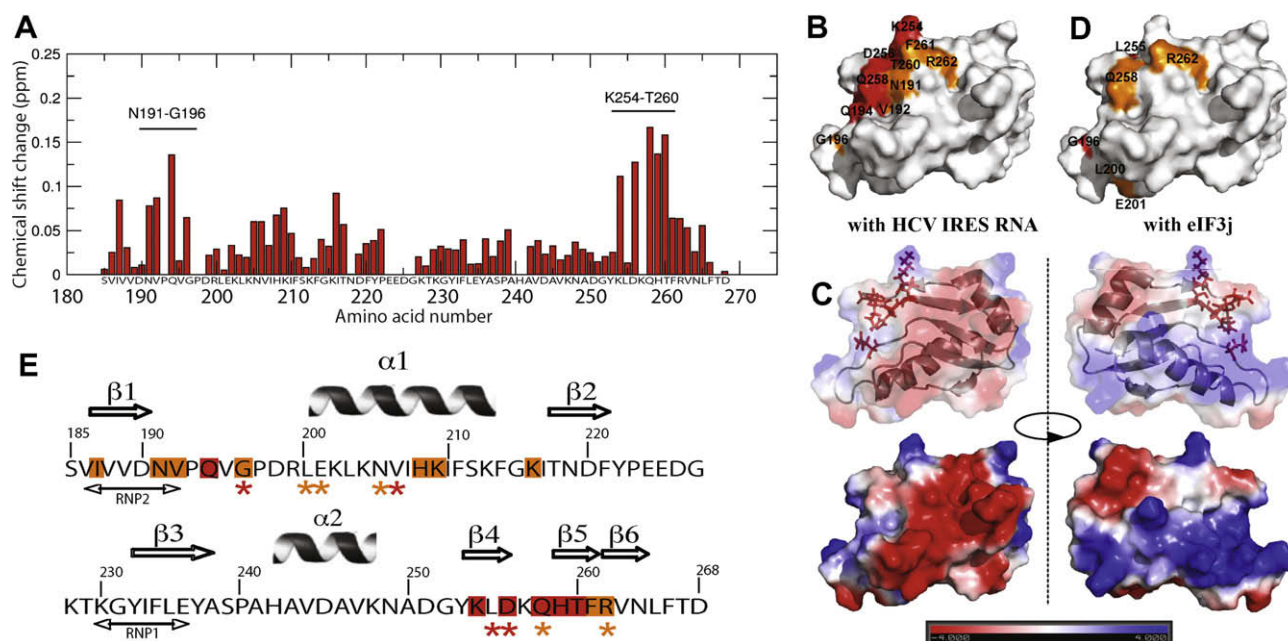


Fig. 4. Mapping eIF3b-RRM/HCV IRES RNA interaction site. (A) Bar graph of the weighed chemical shift differences observed between free eIF3b-RRM and eIF3b-RRM/HCV IRES RNA in 1/1.25 ratio, plotted against the residue number. The two regions showing the largest perturbations upon interaction are indicated. Weighed chemical shift differences were calculated for each signal as follows: $\Delta\delta_{\text{weighed}} = [(\Delta\delta_{\text{H}})^2 + (A515N/6)^2]^{1/2}$. (B) Surface representation of eIF3b-RRM showing the regions of interaction with IRES as determined by NMR titration. Residues having a ppm >0.1 are painted in red, those between 0.06 and 0.1 in orange. (C) Surface potential of eIF3b-RRM. The positively charged residues are indicated in blue and the negatively charged residues in red. Surface area on the opposite side is shown beside. Involved residues in RNA interaction are visible through semi-transparent map. The run was carried out using the PDB structure of eIF3b-RRM using Pymol. (D) Surface representation of eIF3b-RRM showing aa involved in eIF3j-RRM binding (results adapted from Ref. [11]). Orientation is the same as in (B) and (C). (E) Interacting residues between eIF3b-RRM and HCV IRES-RNA are reported on the primary sequence, using the same color code as in (B). On top of the sequence, the defined secondary motifs of the MRR are shown. Stars indicate the residues of eIF3b-RRM involved in eIF3j interaction [11].

3.5. N-terminal recognition motif of eIF3b binds to HCV IRES RNA

To identify the eIF3b-RRM residues involved in the interaction with HCV IRES RNA, we performed NMR titration experiments. Chemical shift variations were monitored using two-dimensional ^1H - ^{15}N correlation spectra. Upon addition of increasing amounts of unlabeled IRES RNA to ^{15}N -labeled eIF3b-RRM, we could follow the chemical shift perturbations. The most affected resonances clustered in two regions, between N191 and G196 and between K254 and T260 (Fig. 4A). The highest chemical shift variations (above 0.1 ppm) were observed for residues Q194, K254, D256, Q258, H259 and T260 (Fig. 4A). Except for residue Q194 located in loop 1, all the residues are located in strands β 4, β 5 and loop5 (Fig. 4E). Moderate chemical shift variations (between 0.06 and 0.1 ppm) were observed for residues I187, N191, V192, G196, H208, H209, K216, F261 and R262, located in RNP2 motif, helix α 1 and strand β 5. Interestingly, all these residues are located on a protruding region separating the negatively charged face from the positively charged face of the RRM (Fig. 4C). This binding mode is different from the consensus RNA-binding surface of RRMs, encompassing the RNP1 and RNP2 motifs (Fig. 4E), but resembles the one observed for PTB RRM3 (polypyrimidine tract binding protein) in complex with RNA, where loop five represents a crucial interaction region for RNA binding (reviewed in [14]).

ElAntack et al. [11] previously showed that the RRM of eIF3b binds to subunit eIF3j (K_D of 20 μM). This interaction involves nine residues, listed in Fig. 4E. Because four of them are also involved in HCV RNA binding, we could think that both partners are exclusive. We found a dissociation constant (K_D) of 2 μM between RNA domain III and eIF3b-RRM, 10 times higher as the one estimated for the eIF3b-RRM/eIF3j-RRM interaction, which would suggest that the IRES element would be favoured (as opposed for eIF3j) in the eIF3 complex binding. Human eIF3j was shown to be required for

the stable binding of eIF3 to 40S ribosomal subunit without other initiation factors. eIF3j is located in the ribosomal decoding center, regulating mRNA binding to the ribosome [15], thus playing an important role in cap-dependent translation initiation. By directly interacting with the RRM of eIF3b, the HCV IRES might influence or even bypass the function of eIF3j that during cap-driven translation initiation associates with the same binding region. Furthermore, our data imply that the HCV IRES functionally compensates the bridging function of eIF3j by independently recruiting the 40S ribosomal subunit and the eIF3 complex.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank Matthias Hentze for critically reading the manuscript, Larissa Balakireva and Patrick Meresse for the monoclonal antibodies. J.P. is the recipient of MENRT fellowship and J.M. of an EMBO fellowship. This work was supported by an ANRS grant 06378-CSS4-2006/2. J.B. was a recipient of a Career Development Award from HFSP (Grant CDA0029/2004-C).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2008.11.025.

References

- [1] Jackson, R.J. and Kaminski, A. (1995) Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA* 1, 985–1000.

- [2] Kieft, J.S., Zhou, K., Jubin, R. and Doudna, J.A. (2001) Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* 7, 194–206.
- [3] Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J. and Hellen, C.U. (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.* 12, 67–83.
- [4] Pisarev, A.V., Shirokikh, N.E. and Hellen, C.U. (2005) Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites. *C.R. Biol.* 328, 589–605.
- [5] Damoc, E., Fraser, C.S., Zhou, M., Videler, H., Mayeur, G.L., Hershey, J.W., Doudna, J.A., Robinson, C.V. and Leary, J.A. (2007) Structural characterization of the human eukaryotic initiation factor 3 protein complex by mass spectrometry. *Mol. Cell Proteom.* 6, 1135–1146.
- [6] Siridechadilok, B., Fraser, C.S., Hall, R.J., Doudna, J.A. and Nogales, E. (2005) Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* 310, 1513–1515.
- [7] Sizova, D.V., Kolupaeva, V.G., Pestova, T.V., Shatsky, I.N. and Hellen, C.U. (1998) Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. *J. Virol.* 72, 4775–4782.
- [8] Buratti, E., Tisminetzky, S., Zotti, M. and Baralle, F.E. (1998) Functional analysis of the interaction between HCV 5'UTR and putative subunits of eukaryotic translation initiation factor eIF3. *Nucl. Acids Res.* 26, 3179–3187.
- [9] Fraser, C.S., Lee, J.Y., Mayeur, G.L., Bushell, M., Doudna, J.A. and Hershey, J.W. (2004) The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40S ribosomal subunits in vitro. *J. Biol. Chem.* 279, 8946–8956.
- [10] Zhou, M., Sandercock, A.M., Fraser, C.S., Ridlova, G., Stephens, E., Schenauer, M.R., Yokoi-Fong, T., Barsky, D., Leary, J.A., Hershey, J.W., Doudna, J.A. and Robinson, C.V. (2008) Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. *Proc. Natl. Acad. Sci. USA.* 105, 18139–18144.
- [11] ElAntak, L., Tzakos, A.G., Locker, N. and Lukavsky, P.J. (2007) Structure of eIF3b RNA recognition motif and its interaction with eIF3j: structural insights into the recruitment of eIF3b to the 40S ribosomal subunit. *J. Biol. Chem.* 282, 8165–8174.
- [12] Pestova, T.V., Hellen, C.U. and Shatsky, I.N. (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell Biol.* 16, 6859–6869.
- [13] Block, K.L., Vornlocher, H.P. and Hershey, J.W. (1998) Characterization of cDNAs encoding the p44 and p35 subunits of human translation initiation factor eIF3. *J. Biol. Chem.* 273, 31901–31908.
- [14] Clery, A., Blatter, M. and Allain, F.H. (2008) RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* 18, 290–298.
- [15] Fraser, C.S., Berry, K.E., Hershey, J.W. and Doudna, J.A. (2007) eIF3j is located in the decoding center of the human 40S ribosomal subunit. *Mol. Cell* 26, 811–819.

SUPPLEMENTARY INFORMATION

Further details of the Methods

S1 : Purification of recombinant human eIF3b-RRM

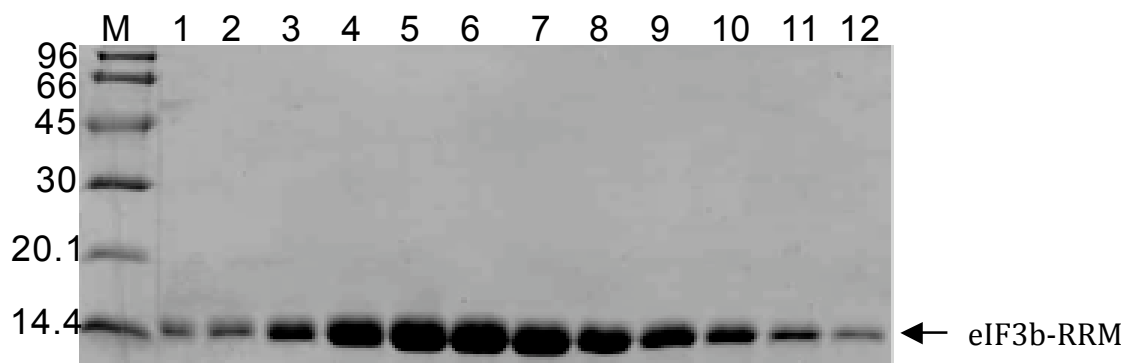
S2 : Purification of eIF3 from rabbit reticulocyte lysate

S3 : NMR labeled protocol

S4 : Fluorescence spectrum of recombinant human eIF3b

S1 : Purification of recombinant human eIF3b-RRM

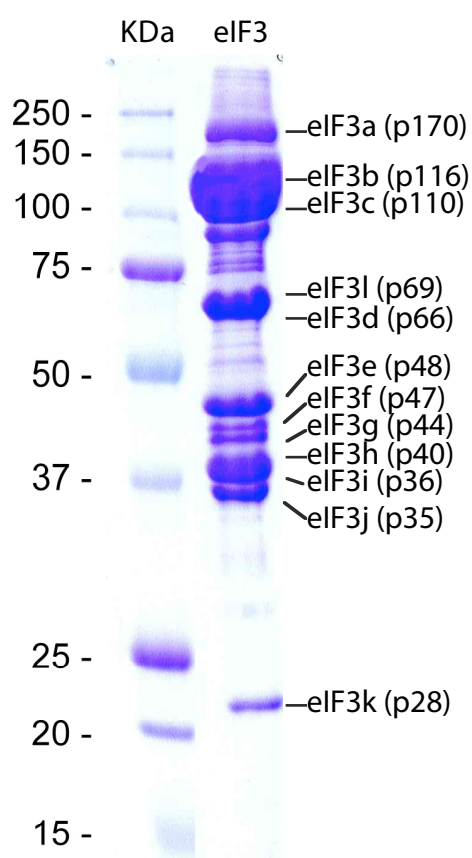
Proteins were expressed in *E. coli* BL21 (DE3) and purified using a His-Select Cobalt column (SIGMA). Lysis buffer : Tris-HCl 20mM pH 7.5, 150mM NaCl, 10mM Imidazole, 3mM Beta-mercapto-ethanol. Washing buffer : Tris-HCl 20mM pH 7.5, 500mM NaCl, 10mM Imidazole, 3mM Beta-mercapto-ethanol. Elution buffer : Tris-HCl 20mM pH 7.5, 150mM NaCl, 300mM Imidazole, 3mM Beta-mercapto-ethanol. Eluted fractions were pooled and concentrated prior to size exclusion chromatography using a Sephadex 75 column (GE Healthcare), in the following buffer : Tris-HCl 20mM pH 7.5, 150mM NaCl, 3mM Beta-mercapto-ethanol. The fractions were collected and concentrated up to 1mM with Amicon ULTRA 5kD concentration unit. Subsequently proteins were analyzed by Laemmli SDS-PAGE 15% (FigS1).

Fig. S1 : Human recombinant eIF3b-RRM protein purification

S2 : Purification of eIF3 from rabbit reticulocyte lysate

The eIF3 complex was purified from rabbit reticulocyte lysate (RRL, Green Hectares) essentially as described in (Pestova et al., 1996 and references therein). Briefly a 0.5M ribosomal salt wash was precipitated with ammonium sulphate (0-40% w/v), collected by centrifugation, and subjected to DEAE- and Phosphocellulose affinity purification. After 10-30% sucrose gradient density centrifugation fractions containing the eIF3 complex were dialyzed against 20mM Tris-HCl pH 7.5, 10% Glycerol, 100mM KCl, 0.1mM EDTA, and 2mM DTT (FigS2).

Fig. S2 : Purification of eIF3 from rabbit reticulocyte lysate



S3 : NMR labeled protocol :

Double labeled ^{13}C , ^{15}N -RRM was produced in *E. coli* using M9 minimal medium supplemented with 1g/L of $^{15}\text{NH}_4\text{Cl}$ and 2g/L [^{13}C] glucose. The protein backbone resonances were assigned by recording tailored triple resonance experiments (Lescop and Brutscher, 2007) ² on a Varian INOVA spectrometers operating at ^1H frequencies of 600 MHz and equipped with a cryogenic triple-resonance (^1H , ^{15}N , ^{13}C) probe with shielded z-gradients. NMR spectra were processed with NMRPipe (Delaglio et al., 1995) ¹, (Fig S3a). Fig S3b represents a region of the ^{15}N - ^1H TROSY spectra of human eIF3b-RRM showing the resonance shifts caused by the interaction with HCV-IRES-RNA, in the following ratio: Black = free eIF3b-RRM. Blue = 4:1 (eIF3b-RRM: RNA in molar ratios) and red, 1:1.25 (eIF3b-RRM: RNA in molar ratios).

Fig. S3a : 2D NMR spectra of eIF3b-RRM and eIF3b-RRM: HCV-IRES-RNA)

A : NMR spectrum of eIF3b-RRM recorded on 800MHz spectrometer

B : NMR spectrum of eIF3b-RRM: HCV-IRES-RNA record on 800MHz spectrometer

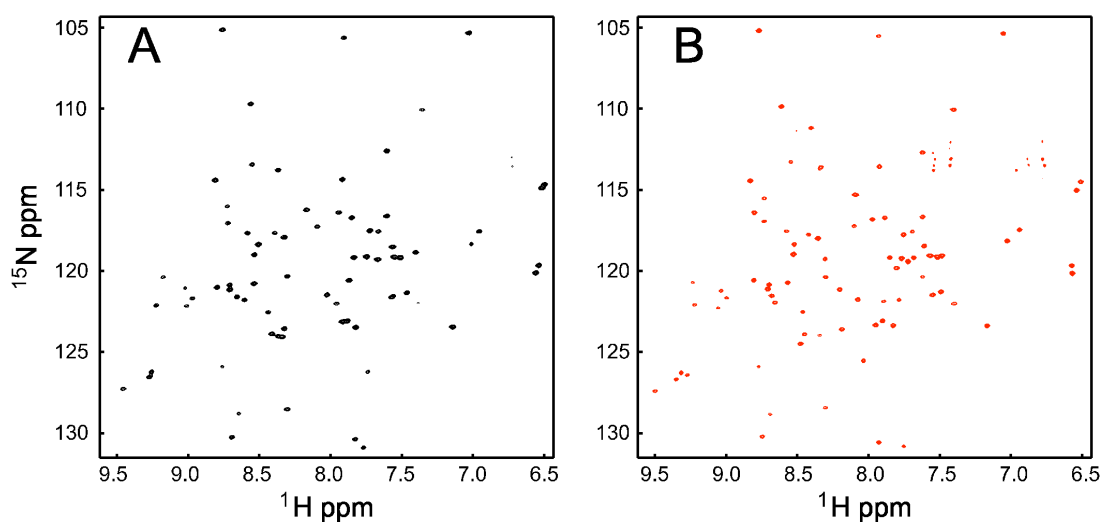
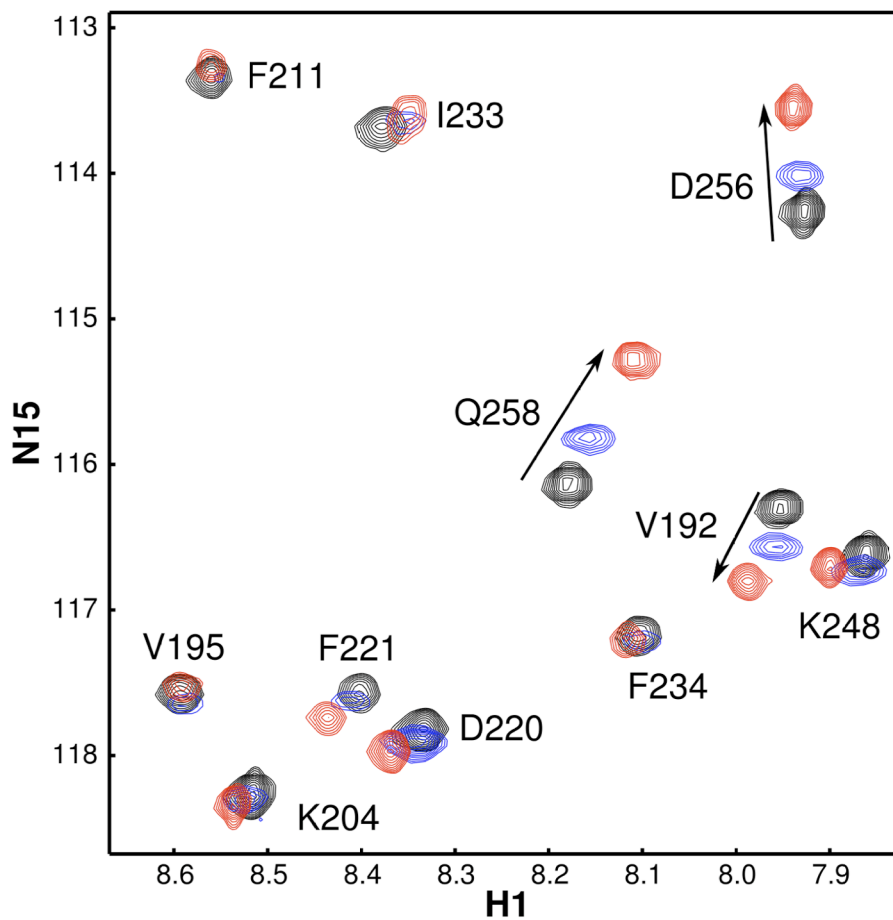


Fig. S3b : Zoom spectra of 2D NMR spectra of eIF3b-RRM and eIF3b-RRM:HCV-IRES-RNA

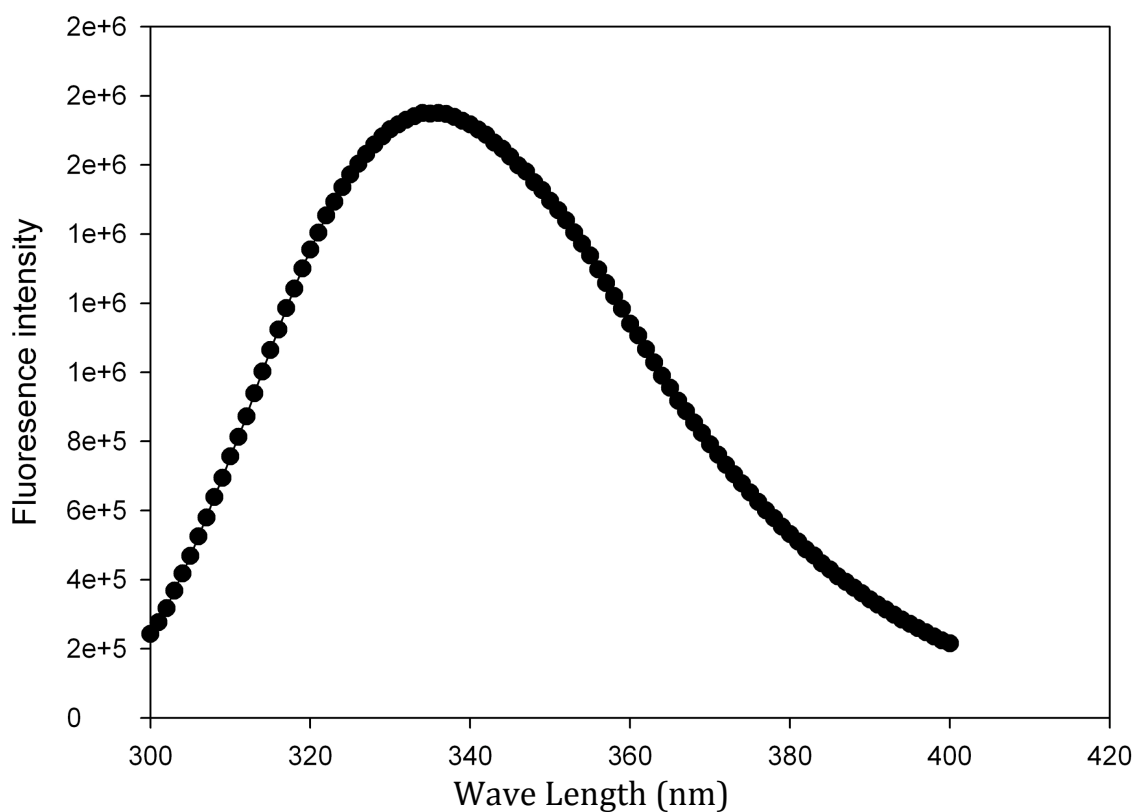


References :

- 1 - Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* 6(3), 277-293.
- 2 - Lescop, E. and Brutscher, B. (2007). Hyperdimensional protein NMR spectroscopy in peptide-sequence space. *J Am Chem Soc* 129(39), 11916-7.

S4 : Fluorescence spectrum of recombinant human eIF3b :

Fluorescence spectrum of recombinant His-Tag eIF3b is recording using 0.1mM of protein. The excitation wavelength is fixed at 288nm and the emission spectra are recorded between 300 and 400nm. The shape of the curve with a maximum emission spectra at 335nm indicated that the protein is properly folded (FigS4).

Fig. S4 : Fluorescence spectrum of recombinant His-Tag eIF3b protein.

REFERENCES BIBLIOGRAPHIQUES

A

Ali, N. and A. Siddiqui (1997). «The La antigen binds 5' noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation.» Proc Natl Acad Sci U S A **94**(6): 2249-54.

Alter, H. J., R. H. Purcell, et al. (1978). «Transmissible agent in non-A, non-B hepatitis.» Lancet **1**(8062): 459-63.

Andre, P., F. Komurian-Pradel, et al. (2002). «Characterization of low- and very-low-density hepatitis C virus RNA-containing particles.» J Virol **76**(14): 6919-28.

Asano, K., T. G. Kinzy, et al. (1997). «Conservation and diversity of eukaryotic translation initiation factor eIF3.» J Biol Chem **272**(2): 1101-9.

Attal, J., M. C. Theron, et al. (1996). «The RU5 ('R') region from human leukaemia viruses (HTLV-1) contains an internal ribosome entry site (IRES)-like sequence.» FEBS Lett **392**(3): 220-4.

B

Balvay, L., R. S. Rifo, et al. (2009). «Structural and functional diversity of viral IRESes.» Biochim Biophys Acta.

Baird, S. D., M. Turcotte, et al. (2006). «Searching for IRES.» Rna **12**(10): 1755-85.

Bandyopadhyay, A. and U. Maitra (1999). «Cloning and characterization of the p42 subunit of mammalian translation initiation factor 3 (eIF3): demonstration that eIF3 interacts with eIF5 in mammalian cells.» Nucleic Acids Res **27**(5): 1331-7.

Bartenschlager, R., L. Ahlborn-Laake, et al. (1993). «Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions.» J Virol **67**(7): 3835-44.

Bartenschlager, R., M. Frese, et al. (2004). «Novel insights into hepatitis C virus replication and persistence.» Adv Virus Res **63**: 71-180.

Baumert, T. F., S. Ito, et al. (1998). «Hepatitis C virus structural proteins assemble into viruslike particles in insect cells.» J Virol **72**(5): 3827-36.

Beales, L. P., D. J. Rowlands, et al. (2001). «The internal ribosome entry site (IRES) of hepatitis C virus visualized by electron microscopy.» Rna **7**(5): 661-70.

Behrens, S. E., L. Tomei, et al. (1996). «Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus.» Embo J **15**(1): 12-22.

Belsham, G. J. (1992). «Dual initiation sites of protein synthesis on foot-and-mouth disease virus RNA are selected following internal entry and scanning of ribosomes in vivo.» Embo J **11**(3): 1105-10.

Belsham, G. J. (2009). «Divergent picornavirus IRES elements.» *Virus Res* **139**(2): 183-92.

Belsham, G. J. and J. K. Brangwyn (1990). «A region of the 5' noncoding region of foot-and-mouth disease virus RNA directs efficient internal initiation of protein synthesis within cells: involvement with the role of L protease in translational control.» *J Virol* **64**(11): 5389-95.

Blumberg, B. S., L. Melartin, et al. (1966). «Family studies of a human serum isoantigen system (Australia antigen).» *Am J Hum Genet* **18**(6): 594-608.

Boehringer, D., R. Thermann, et al. (2005). «Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES.» *Structure* **13**(11): 1695-706.

Borman, A., M. T. Howell, et al. (1993). «The involvement of a spliceosome component in internal initiation of human rhinovirus RNA translation.» *J Gen Virol* **74** (Pt 9): 1775-88.

Borman, A. and R. J. Jackson (1992). «Initiation of translation of human rhinovirus RNA: mapping the internal ribosome entry site.» *Virology* **188**(2): 685-96.

Brown, E. A., H. Zhang, et al. (1992). «Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs.» *Nucleic Acids Res* **20**(19): 5041-5.

Browning, K. S., D. R. Gallie, et al. (2001). «Unified nomenclature for the subunits of eukaryotic initiation factor 3.» *Trends Biochem Sci* **26**(5): 284.

Buck, C. B., X. Shen, et al. (2001). «The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site.» *J Virol* **75**(1): 181-91.

Byrd, M. P., M. Zamora, et al. (2005). «Translation of eukaryotic translation initiation factor 4GI (eIF4GI) proceeds from multiple mRNAs containing a novel cap-dependent internal ribosome entry site (IRES) that is active during poliovirus infection.» *J Biol Chem* **280**(19): 18610-22.

C

Carter, M. S. and P. Sarnow (2000). «Distinct mRNAs that encode La autoantigen are differentially expressed and contain internal ribosome entry sites.» *J Biol Chem* **275**(36): 28301-7.

Chaudhuri, J., A. Chakrabarti, et al. (1997). «Biochemical characterization of mammalian translation initiation factor 3 (eIF3). Molecular cloning reveals that p110 subunit is the mammalian homologue of *Saccharomyces cerevisiae* protein Prt1.» *J Biol Chem* **272**(49): 30975-83.

Chevaliez, S. and J. M. Pawlotsky (2008). «Diagnosis and management of chronic viral hepatitis: antigens, antibodies and viral genomes.» *Best Pract Res Clin Gastroenterol* **22**(6): 1031-48.

Choo, Q. L., G. Kuo, et al. (1989). «Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome.» *Science* **244**(4902): 359-62.

Cocquerel, L., S. Duvet, et al. (1999). «The transmembrane domain of hepatitis C virus glycoprotein E1 is a signal for static retention in the endoplasmic reticulum.» *J Virol* **73**(4):

2641-9.

Cornelis, S., Y. Bruynooghe, et al. (2000). «Identification and characterization of a novel cell cycle-regulated internal ribosome entry site.» *Mol Cell* **5**(4): 597-605.

Correll, C. C., B. Freeborn, et al. (1997). «Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain.» *Cell* **91**(5): 705-12.

D

Damoc, E., C. S. Fraser, et al. (2007). «Structural characterization of the human eukaryotic initiation factor 3 protein complex by mass spectrometry.» *Mol Cell Proteomics* **6**(7): 1135-46.

Dane, D. S., C. H. Cameron, et al. (1970). «Virus-like particles in serum of patients with Australia-antigen-associated hepatitis.» *Lancet* **1**(7649): 695-8.

Deffaud, C. and J. L. Darlix (2000). «Characterization of an internal ribosomal entry segment in the 5' leader of murine leukemia virus env RNA.» *J Virol* **74**(2): 846-50.

Deforges, S., A. Evlashev, et al. (2004). «Expression of hepatitis C virus proteins in epithelial intestinal cells in vivo.» *J Gen Virol* **85**(Pt 9): 2515-23.

E

ElAntak, L., A. G. Tzacos, et al. (2007). «Structure of eIF3b RNA recognition motif and its interaction with eIF3j: structural insights into the recruitment of eIF3b to the 40 S ribosomal subunit.» *J Biol Chem* **282**(11): 8165-74.

Ezelle, H. J., D. Markovic, et al. (2002). «Generation of hepatitis C virus-like particles by use of a recombinant vesicular stomatitis virus vector.» *J Virol* **76**(23): 12325-34.

F

Feinstone, S. M., A. Z. Kapikian, et al. (1973). «Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness.» *Science* **182**(116): 1026-8.

Feinstone, S. M., A. Z. Kapikian, et al. (1975). «Transfusion-associated hepatitis not due to viral hepatitis type A or B.» *N Engl J Med* **292**(15): 767-70.

Feinstone, S. M., K. B. Mihalik, et al. (1983). «Inactivation of hepatitis B virus and non-A, non-B hepatitis by chloroform.» *Infect Immun* **41**(2): 816-21.

Feld, J. J. and T. J. Liang (2005). «HCV persistence: cure is still a four letter word.» *Hepatology* **41**(1): 23-5.

Fernandez, J., I. Yaman, et al. (2001). «Internal ribosome entry site-mediated translation of a mammalian mRNA is regulated by amino acid availability.» *J Biol Chem* **276**(15): 12285-91.

Fernandez-Miragall, O., S. Lopez de Quinto, et al. (2009). «Relevance of RNA structure for the

activity of picornavirus IRES elements.» Virus Res **139**(2): 172-82.

Fletcher, C. M., T. V. Pestova, et al. (1999). «Structure and interactions of the translation initiation factor eIF1.» Embo J **18**(9): 2631-7.

Forton, D. M., P. Karayiannis, et al. (2004). «Identification of unique hepatitis C virus quasispecies in the central nervous system and comparative analysis of internal translational efficiency of brain, liver, and serum variants.» J Virol **78**(10): 5170-83.

Fraser, C. S., K. E. Berry, et al. (2007). «eIF3j is located in the decoding center of the human 40S ribosomal subunit.» Mol Cell **26**(6): 811-9.

Fraser, C. S. and J. A. Doudna (2007). «Structural and mechanistic insights into hepatitis C viral translation initiation.» Nat Rev Microbiol **5**(1): 29-38.

Fraser, C. S., J. W. Hershey, et al. (2009). «The pathway of hepatitis C virus mRNA recruitment to the human ribosome.» Nat Struct Mol Biol **16**(4): 397-404.

Fraser, C. S., J. Y. Lee, et al. (2004). «The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40 S ribosomal subunits in vitro.» J Biol Chem **279**(10): 8946-56.

Fried, M. W. (2002). «Side effects of therapy of hepatitis C and their management.» Hepatology **36**(5 Suppl 1): S237-44.

Fukushi, S., M. Okada, et al. (2001). «Ribosomal protein S5 interacts with the internal ribosomal entry site of hepatitis C virus.» J Biol Chem **276**(24): 20824-6.

G

Gallie, D. R. (2001). «Cap-independent translation conferred by the 5' leader of tobacco etch virus is eukaryotic initiation factor 4G dependent.» J Virol **75**(24): 12141-52.

Glass, M. J., X. Y. Jia, et al. (1993). «Identification of the hepatitis A virus internal ribosome entry site: in vivo and in vitro analysis of bicistronic RNAs containing the HAV 5' noncoding region.» Virology **193**(2): 842-52.

Gosert, R., D. Egger, et al. (2003). «Identification of the hepatitis C virus RNA replication complex in Huh-7 cells harboring subgenomic replicons.» J Virol **77**(9): 5487-92.

Graff, J., J. Cha, et al. (1998). «Interaction of poly(rC) binding protein 2 with the 5' noncoding region of hepatitis A virus RNA and its effects on translation.» J Virol **72**(12): 9668-75.

Grakoui, A., D. W. McCourt, et al. (1993a). «A second hepatitis C virus-encoded proteinase.» Proc Natl Acad Sci U S A **90**(22): 10583-7.

Grakoui, A., D. W. McCourt, et al. (1993b). «Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites.» J Virol **67**(5): 2832-43.

- Grakoui, A., C. Wychowski, et al. (1993c). «Expression and identification of hepatitis C virus polyprotein cleavage products.» *J Virol* **67**(3): 1385-95.
- Grundhoff, A. and D. Ganem (2001). «Mechanisms governing expression of the v-FLIP gene of Kaposi's sarcoma-associated herpesvirus.» *J Virol* **75**(4): 1857-63.
- Guha, C., S. W. Lee, et al. (2005). «Cell culture models and animal models of viral hepatitis. Part II: hepatitis C.» *Lab Anim (NY)* **34**(2): 39-47.
- Gunjan, A., J. Paik, et al. (2005). «Regulation of histone synthesis and nucleosome assembly.» *Biochimie* **87**(7): 625-35.

H

- Hahm, B., Y. K. Kim, et al. (1998). «Heterogeneous nuclear ribonucleoprotein L interacts with the 3' border of the internal ribosomal entry site of hepatitis C virus.» *J Virol* **72**(11): 8782-8.
- Hellen, C. U. and S. de Breyne (2007). «A distinct group of hepacivirus/pestivirus-like internal ribosomal entry sites in members of diverse picornavirus genera: evidence for modular exchange of functional noncoding RNA elements by recombination.» *J Virol* **81**(11): 5850-63.
- Hellen, C. U., T. V. Pestova, et al. (1994). «The cellular polypeptide p57 (pyrimidine tract-binding protein) binds to multiple sites in the poliovirus 5' nontranslated region.» *J Virol* **68**(2): 941-50.
- Henis-Korenblit, S., N. L. Strumpf, et al. (2000). «A novel form of DAP5 protein accumulates in apoptotic cells as a result of caspase cleavage and internal ribosome entry site-mediated translation.» *Mol Cell Biol* **20**(2): 496-506.
- Hernandez, G., P. Vazquez-Pianzola, et al. (2004). «Internal ribosome entry site drives cap-independent translation of reaper and heat shock protein 70 mRNAs in *Drosophila* embryos.» *Rna* **10**(11): 1783-97.
- Hijikata, M., N. Kato, et al. (1991). «Hypervariable regions in the putative glycoprotein of hepatitis C virus.» *Biochem Biophys Res Commun* **175**(1): 220-8.
- Holcik, M. (2004). «Targeting translation for treatment of cancer--a novel role for IRES?» *Curr Cancer Drug Targets* **4**(3): 299-311.
- Holcik, M. and N. Sonenberg (2005). «Translational control in stress and apoptosis.» *Nat Rev Mol Cell Biol* **6**(4): 318-27.
- Holcik, M., N. Sonenberg, et al. (2000). «Internal ribosome initiation of translation and the control of cell death.» *Trends Genet* **16**(10): 469-73.
- Honda, M., E. A. Brown, et al. (1996a). «Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA.» *Rna* **2**(10): 955-68.

Honda, M., L. H. Ping, et al. (1996b). «Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA.» *Virology* **222**(1): 31-42.

Honda, M., R. Rijnbrand, et al. (1999). «Natural variation in translational activities of the 5' nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long-range RNA-RNA interaction outside of the internal ribosomal entry site.» *J Virol* **73**(6): 4941-51.

Hunt, S. L., J. J. Hsuan, et al. (1999). «unr, a cellular cytoplasmic RNA-binding protein with five cold-shock domains, is required for internal initiation of translation of human rhinovirus RNA.» *Genes Dev* **13**(4): 437-48.

I

Ishida, S., M. Kaito, et al. (2001). «Hepatitis C virus core particle detected by immunoelectron microscopy and optical rotation technique.» *Hepato Res* **20**(3): 335-347.

J

Jackson, R. J., S. L. Hunt, et al. (1994). «Internal initiation of translation of picornavirus RNAs.» *Mol Biol Rep* **19**(3): 147-59.

Jackson, R. J. and A. Kaminski (1995). «Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond.» *Rna* **1**(10): 985-1000.

Jang, C. J., M. C. Lo, et al. (2009). «Conserved element of the dicistrovirus IGR IRES that mimics an E-site tRNA/ribosome interaction mediates multiple functions.» *J Mol Biol* **387**(1): 42-58.

Jang, S. K., H. G. Krausslich, et al. (1988). «A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation.» *J Virol* **62**(8): 2636-43.

Ji, H., C. S. Fraser, et al. (2004). «Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA.» *Proc Natl Acad Sci U S A* **101**(49): 16990-5.

Johannes, G. and P. Sarnow (1998). «Cap-independent polysomal association of natural mRNAs encoding c-myc, BiP, and eIF4G conferred by internal ribosome entry sites.» *Rna* **4**(12): 1500-13.

Jopling, C. L., M. Yi, et al. (2005). «Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA.» *Science* **309**(5740): 1577-81.

K

Kaito, M., S. Watanabe, et al. (1994). «Hepatitis C virus particle detected by immunoelectron microscopic study.» *J Gen Virol* **75 (Pt 7)**: 1755-60.

- Kapp, L. D. and J. R. Lorsch (2004). «The molecular mechanics of eukaryotic translation.» Annu Rev Biochem **73**: 657-704.
- Kato, T., T. Date, et al. (2003). «Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon.» Gastroenterology **125**(6): 1808-17.
- Kieft, J. S. (2008). «Viral IRES RNA structures and ribosome interactions.» Trends Biochem Sci **33**(6): 274-83.
- Kieft, J. S., K. Zhou, et al. (2002). «Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation.» Nat Struct Biol **9**(5): 370-4.
- Kieft, J. S., K. Zhou, et al. (1999). «The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold.» J Mol Biol **292**(3): 513-29.
- Koch, M. H., P. Vachette, et al. (2003). «Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution.» Q Rev Biophys **36**(2): 147-227.
- Kolupaeva, V. G., S. de Breyne, et al. (2007). «In vitro reconstitution and biochemical characterization of translation initiation by internal ribosomal entry.» Methods Enzymol **430**: 409-39.
- Kolupaeva, V. G., C. U. Hellen, et al. (1996). «Structural analysis of the interaction of the pyrimidine tract-binding protein with the internal ribosomal entry site of encephalomyocarditis virus and foot-and-mouth disease virus RNAs.» Rna **2**(12): 1199-212.
- Kolupaeva, V. G., T. V. Pestova, et al. (2000a). «An enzymatic footprinting analysis of the interaction of 40S ribosomal subunits with the internal ribosomal entry site of hepatitis C virus.» J Virol **74**(14): 6242-50.
- Kolupaeva, V. G., T. V. Pestova, et al. (2000b). «Ribosomal binding to the internal ribosomal entry site of classical swine fever virus.» Rna **6**(12): 1791-807.
- Kolupaeva, V. G., A. Unbehaun, et al. (2005). «Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association.» Rna **11**(4): 470-86.
- Koshy, A., J. P. Madda, et al. (2002). «Treatment of hepatitis C virus genotype 4-related cirrhosis: ribavirin and interferon combination compared with interferon alone.» J Clin Gastroenterol **35**(1): 82-5.
- Kozak, M. (1980). «Evaluation of the «scanning model» for initiation of protein synthesis in eucaryotes.» Cell **22**(1 Pt 1): 7-8.
- Kozak, M. (1997). «Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6.» EMBO J **16**(9): 2482-92.

Kuo, G., Q. L. Choo, et al. (1989). «An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis.» Science **244**(4902): 362-4.

L

Lancaster, A. M., E. Jan, et al. (2006). «Initiation factor-independent translation mediated by the hepatitis C virus internal ribosome entry site.» Rna **12**(5): 894-902.

Layden-Almer, J. E., R. M. Ribeiro, et al. (2003). «Viral dynamics and response differences in HCV-infected African American and white patients treated with IFN and ribavirin.» Hepatology **37**(6): 1343-50.

Le, S. Y., W. M. Liu, et al. (1998). «Phylogenetic evidence for the improved RNA higher-order structure in internal ribosome entry sequences of HCV and pestiviruses.» Virus Genes **17**(3): 279-95.

Lin, H. H., H. Y. Hsu, et al. (1994). «Hepatitis C virus infection in pregnant women: detection by different anti-HCV immunoassays and serum HCV-RNA.» Asia Oceania J Obstet Gynaecol **20**(1): 13-8.

Locker, N., L. E. Easton, et al. (2007). «HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly.» EMBO J **26**(3): 795-805.

Lohmann, V., F. Korner, et al. (1999). «Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line.» Science **285**(5424): 110-3.

Lukavsky, P. J. (2008). «Structure and function of HCV IRES domains.» Virus Res.

Mac, C. F. (1947). «Collection of specimens for virus investigations.» Mon Bull Minist Health Public Health Lab Serv **6**: 229-32.

M

Macejak, D. G. and P. Sarnow (1991). «Internal initiation of translation mediated by the 5' leader of a cellular mRNA.» Nature **353**(6339): 90-4.

Maillard, P., K. Krawczynski, et al. (2001). «Nonenveloped nucleocapsids of hepatitis C virus in the serum of infected patients.» J Virol **75**(17): 8240-50.

Majumdar, A., A. Dasgupta, et al. (1979). «Purification and properties of rabbit reticulocyte protein synthesis initiation factors EIF-1, EIF-2, and EIF-3.» Methods Enzymol **60**: 35-52.

Majumdar, R., A. Bandyopadhyay, et al. (2003). «Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex.» J Biol Chem **278**(8): 6580-7.

Malnou, C. E., T. A. Poyry, et al. (2002). «Poliovirus internal ribosome entry segment structure alterations that specifically affect function in neuronal cells: molecular genetic analysis.» J Virol **76**(21): 10617-26.

Manns, M. P. and C. P. Strassburg (2001). «Autoimmune hepatitis: clinical challenges.» Gastroenterology **120**(6): 1502-17.

Marash, L. and A. Kimchi (2005). «DAP5 and IRES-mediated translation during programmed cell death.» Cell Death Differ **12**(6): 554-62.

Martinez-Salas, E., R. Ramos, et al. (2001). «Functional interactions in internal translation initiation directed by viral and cellular IRES elements.» J Gen Virol **82**(Pt 5): 973-84.

Masutani, M., N. Sonenberg, et al. (2007). «Reconstitution reveals the functional core of mammalian eIF3.» Embo J **26**(14): 3373-83.

Mayeur, G. L., C. S. Fraser, et al. (2003). «Characterization of eIF3k: a newly discovered subunit of mammalian translation initiation factor eIF3.» Eur J Biochem **270**(20): 4133-9.

Meerovitch, K., J. Pelletier, et al. (1989). «A cellular protein that binds to the 5'-noncoding region of poliovirus RNA: implications for internal translation initiation.» Genes Dev **3**(7): 1026-34.

Melcher, S. E., T. J. Wilson, et al. (2003). «The dynamic nature of the four-way junction of the hepatitis C virus IRES.» Rna **9**(7): 809-20.

Methot, N., E. Rom, et al. (1997). «The human homologue of the yeast Prt1 protein is an integral part of the eukaryotic initiation factor 3 complex and interacts with p170.» J Biol Chem **272**(2): 1110-6.

Methot, N., M. S. Song, et al. (1996). «A region rich in aspartic acid, arginine, tyrosine, and glycine (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-association and interaction with eIF3.» Mol Cell Biol **16**(10): 5328-34.

Mokrejs, M., V. Vopalensky, et al. (2006). «IRESite: the database of experimentally verified IRES structures (www.iresite.org).» Nucleic Acids Res **34**(Database issue): D125-30.

Moradpour, D., R. Gosert, et al. (2003). «Membrane association of hepatitis C virus nonstructural proteins and identification of the membrane alteration that harbors the viral replication complex.» Antiviral Res **60**(2): 103-9.

N

Nanbru, C., I. Lafon, et al. (1997). «Alternative translation of the proto-oncogene c-myc by an internal ribosome entry site.» J Biol Chem **272**(51): 32061-6.

Nothdurft, H. D. (2008). «Hepatitis A vaccines.» Expert Rev Vaccines **7**(5): 535-45.

O

Oh, S. K., M. P. Scott, et al. (1992). «Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding.» Genes Dev **6**(9):

1643-53.

Ohlmann, T., M. Lopez-Lastra, et al. (2000). «An internal ribosome entry segment promotes translation of the simian immunodeficiency virus genomic RNA.» *J Biol Chem* **275**(16): 11899-906.

Ohlmann, T., V. M. Pain, et al. (1997). «The proteolytic cleavage of eukaryotic initiation factor (eIF) 4G is prevented by eIF4E binding protein (PHAS-I; 4E-BP1) in the reticulocyte lysate.» *EMBO J* **16**(4): 844-55.

Ohnuma, H., T. Tanaka, et al. (2001). «The first large-scale nucleic acid amplification testing (NAT) of donated blood using multiplex reagent for simultaneous detection of HBV, HCV, and HIV-1 and significance of NAT for HBV.» *Microbiol Immunol* **45**(9): 667-72.

Otto, G. A. and J. D. Puglisi (2004). «The pathway of HCV IRES-mediated translation initiation.» *Cell* **119**(3): 369-80.

P

Pace, C. N., F. Vajdos, et al. (1995). «How to measure and predict the molar absorption coefficient of a protein.» *Protein Sci* **4**(11): 2411-23.

Pelletier, J. and N. Sonenberg (1988). «Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA.» *Nature* **334**(6180): 320-5.

Perez, I., C. H. Lin, et al. (1997). «Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo.» *Rna* **3**(7): 764-78.

Pestova, T. V. and C. U. Hellen (2000). «The structure and function of initiation factors in eukaryotic protein synthesis.» *Cell Mol Life Sci* **57**(4): 651-74.

Pestova, T. V. and C. U. Hellen (2001). «Functions of eukaryotic factors in initiation of translation.» *Cold Spring Harb Symp Quant Biol* **66**: 389-96.

Pestova, T. V. and C. U. Hellen (2003). «Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA.» *Genes Dev* **17**(2): 181-6.

Pestova, T. V., C. U. Hellen, et al. (1991). «Translation of poliovirus RNA: role of an essential cis-acting oligopyrimidine element within the 5' nontranslated region and involvement of a cellular 57-kilodalton protein.» *J Virol* **65**(11): 6194-204.

Pestova, T. V., I. N. Shatsky, et al. (1998). «A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs.» *Genes Dev* **12**(1): 67-83.

Petit, M. A., M. Lievre, et al. (2005). «Enveloped particles in the serum of chronic hepatitis C patients.» *Virology* **336**(2): 144-53.

- Pilipenko, E. V., T. V. Pestova, et al. (2000). «A cell cycle-dependent protein serves as a template-specific translation initiation factor.» *Genes Dev* **14**(16): 2028-45.
- Pinkstaff, J. K., S. A. Chappell, et al. (2001). «Internal initiation of translation of five dendritically localized neuronal mRNAs.» *Proc Natl Acad Sci U S A* **98**(5): 2770-5.
- Pisarev, A. V., N. E. Shirokikh, et al. (2005). «Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites.» *C R Biol* **328**(7): 589-605.
- Poole, T. L., C. Wang, et al. (1995). «Pestivirus translation initiation occurs by internal ribosome entry.» *Virology* **206**(1): 750-4.
- Prats, A. C., S. Vagner, et al. (1992). «cis-acting elements involved in the alternative translation initiation process of human basic fibroblast growth factor mRNA.» *Mol Cell Biol* **12**(10): 4796-805.
- Prevot, D., J. L. Darlix, et al. (2003). «Conducting the initiation of protein synthesis: the role of eIF4G.» *Biol Cell* **95**(3-4): 141-56.
- Psaridi, L., U. Georgopoulou, et al. (1999). «Mutational analysis of a conserved tetraloop in the 5' untranslated region of hepatitis C virus identifies a novel RNA element essential for the internal ribosome entry site function.» *FEBS Lett* **453**(1-2): 49-53.
- Putnam, C. D., M. Hammel, et al. (2007). «X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.» *Q Rev Biophys* **40**(3): 191-285.
- Pybus, O. G., P. V. Markov, et al. (2007). «Investigating the endemic transmission of the hepatitis C virus.» *Int J Parasitol* **37**(8-9): 839-49.

Q

- Qin, X. and P. Sarnow (2004). «Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells.» *J Biol Chem* **279**(14): 13721-8.

R

- Reynolds, J. E., A. Kaminski, et al. (1995). «Unique features of internal initiation of hepatitis C virus RNA translation.» *Embo J* **14**(23): 6010-20.
- Rijnbrand, R., T. van der Straaten, et al. (1997). «Internal entry of ribosomes is directed by the 5' noncoding region of classical swine fever virus and is dependent on the presence of an RNA pseudoknot upstream of the initiation codon.» *J Virol* **71**(1): 451-7.
- Rijnbrand, R. C., T. E. Abbink, et al. (1996). «The influence of AUG codons in the hepatitis C virus 5' nontranslated region on translation and mapping of the translation initiation window.» *Virology* **226**(1): 47-56.

Rijnbrand, R. C. and S. M. Lemon (2000). «Internal ribosome entry site-mediated translation in hepatitis C virus replication.» *Curr Top Microbiol Immunol* **242**: 85-116.

S

Saito, S., N. Kato, et al. (1995). «[A comparative study of hypervariable regions in positive- and negative-stranded HCV-RNAs from hepatocellular carcinoma and the surrounding liver].» *Nippon Rinsho* **53 Suppl**(Pt 1): 800-4.

Santolini, E., G. Migliaccio, et al. (1994). «Biosynthesis and biochemical properties of the hepatitis C virus core protein.» *J Virol* **68**(6): 3631-41.

Selby, M. J., Q. L. Choo, et al. (1993). «Expression, identification and subcellular localization of the proteins encoded by the hepatitis C viral genome.» *J Gen Virol* **74 (Pt 6)**: 1103-13.

Shimizu, Y. K., S. M. Feinstone, et al. (1979). «Non-A, non-B hepatitis: ultrastructural evidence for two agents in experimentally infected chimpanzees.» *Science* **205**(4402): 197-200.

Shine, J. and L. Dalgarno (1975). «Determinant of cistron specificity in bacterial ribosomes.» *Nature* **254**(5495): 34-8.

Simmonds, P. (1995). «Variability of hepatitis C virus.» *Hepatology* **21**(2): 570-83.

Simmonds, P., J. Bukh, et al. (2005). «Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes.» *Hepatology* **42**(4): 962-73.

Siridechadilok, B., C. S. Fraser, et al. (2005). «Structural roles for human translation factor eIF3 in initiation of protein synthesis.» *Science* **310**(5753): 1513-5.

Sizova, D. V., V. G. Kolupaeva, et al. (1998). «Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs.» *J Virol* **72**(6): 4775-82.

Spahn, C. M., J. S. Kieft, et al. (2001). «Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit.» *Science* **291**(5510): 1959-62.

Spangberg, K. and S. Schwartz (1999). «Poly(C)-binding protein interacts with the hepatitis C virus 5' untranslated region.» *J Gen Virol* **80 (Pt 6)**: 1371-6.

Stoneley, M. and A. E. Willis (2004). «Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression.» *Oncogene* **23**(18): 3200-7.

Svitkin, Y. V., K. Meerovitch, et al. (1994). «Internal translation initiation on poliovirus RNA: further characterization of La function in poliovirus translation in vitro.» *J Virol* **68**(3): 1544-50.

T

Tai, C. L., W. K. Chi, et al. (1996). «The helicase activity associated with hepatitis C virus

nonstructural protein 3 (NS3).» *J Virol* **70**(12): 8477-84.

Takahashi, K., H. Okamoto, et al. (1992). «Demonstration of a hepatitis C virus-specific antigen predicted from the putative core gene in the circulation of infected hosts.» *J Gen Virol* **73 (Pt 3)**: 667-72.

Touriol, C., S. Bornes, et al. (2003). «Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons.» *Biol Cell* **95**(3-4): 169-78.

Toyoda, H., H. Sakamoto, et al. (2000). «Eradication of hepatitis C virus 1b by interferon in a health care worker with acute hepatitis following needlestick transmission from a patient with chronic hepatitis C unresponsive to interferon.» *Scand J Gastroenterol* **35**(10): 1117-20.

Tsukiyama-Kohara, K., N. Iizuka, et al. (1992). «Internal ribosome entry site within hepatitis C virus RNA.» *J Virol* **66**(3): 1476-83.

V

Vagner, S., M. C. Gensac, et al. (1995). «Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes.» *Mol Cell Biol* **15**(1): 35-44.

Van Eden, M. E., M. P. Byrd, et al. (2004). «Translation of cellular inhibitor of apoptosis protein 1 (c-IAP1) mRNA is IRES mediated and regulated during cell stress.» *RNA* **10**(3): 469-81.

Voisset, C. and J. Dubuisson (2004). «Functional hepatitis C virus envelope glycoproteins.» *Biol Cell* **96**(6): 413-20.

W

Watanabe, H., T. Saito, et al. (2003). «Spontaneous elimination of serum hepatitis C virus (HCV) RNA in chronic HCV carriers: a population-based cohort study.» *J Med Virol* **71**(1): 56-61.

Wiegand, J., P. Buggisch, et al. (2006). «Early monotherapy with pegylated interferon alpha-2b for acute hepatitis C infection: the HEP-NET acute-HCV-II study.» *Hepatology* **43**(2): 250-6.

Y

Yang, W., Y. Sun, et al. (2007). «Hepatitis C virus (HCV) NS5B nonnucleoside inhibitors specifically block single-stranded viral RNA synthesis catalyzed by HCV replication complexes in vitro.» *Antimicrob Agents Chemother* **51**(1): 338-42.

Ye, X., P. Fong, et al. (1997). «Ultrabithorax and Antennapedia 5' untranslated regions promote developmentally regulated internal translation initiation.» *Mol Cell Biol* **17**(3): 1714-21.

Yu, X., M. Qiao, et al. (2007). «Cryo-electron microscopy and three-dimensional reconstructions of hepatitis C virus particles.» *Virology* **367**(1): 126-34.

Z

Zein, N. N. (2000). «Clinical significance of hepatitis C virus genotypes.» Clin Microbiol Rev **13**(2): 223-35.

Zhao, X., S. McKillop-Smith, et al. (2004). «The human histone gene expression regulator HBP/SLBP is required for histone and DNA synthesis, cell cycle progression and cell proliferation in mitotic cells.» J Cell Sci **117**(Pt 25): 6043-51.

Zhou, M., A. M. Sandercock, et al. (2008). «Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3.» Proc Natl Acad Sci U S A **105**(47): 18139-44.

Zhou, W., G. M. Edelman, et al. (2001). «Transcript leader regions of two *Saccharomyces cerevisiae* mRNAs contain internal ribosome entry sites that function in living cells.» Proc Natl Acad Sci U S A **98**(4): 1531-6.

Livres :

Hershey, J., Merrick, W. The pathway and mechanism of initiation of protein synthesis. In: Sonenberg N., et al., editors. *Translational control of gene expression*. 2nd ed. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2000. pp. 33–88.

Merrick, W., Nyborg, J. The protein biosynthesis elongation cycle. In: Sonenberg N., et al., editors. *Translational control of gene expression*. 2nd ed. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2000. pp. 89–126.

ANNEXES

TABLE DES ANNEXES

ANNEXE I : Tableau détaillé des thérapies anti-VHC disponibles

ANNEXE II : Liste des oligonucléotides générés pour le clonage

Annexe III : Protocole de purification du complexe eIF3 à partir de réticulocytes de lapin.

Annexe IV : Protocole expérimentale utilisé pour le test d'expression à haut débit des différentes constructions des sous-unités du complexe eIF3.

Annexe V : Protocole expérimentale présentant les différentes phases de la technique de SAXS (Svergun).

Annexe I : Tableau détaillé des thérapies anti-VHC disponibles

Cible	Mécanisme	Compagnie pharmaceutique	Phase
IRES	ARN Antisens Ribozyme	ISIS Pharmaceuticals	II
		Elan corporation	III
		Ribozyme Pharmaceuticals	
Protéase NS3/ NS4A	Inhibiteurs NS3/4A	Boehringer ingelheim	III
		Vertex Pharmaceuticals	Pré-clinique
		Boehringer (BILN 2061)	ARRET
		VX 950 (Telaprevir)	III
		SCH 503034	III
		Boceprevir	II
		TMC435	II
		SCH900518	II
		ACH 806/GS 9132	ARRET
ITMN 191	I		
NS5A	Fixation sur NS5A marche sur les réplicons	Arrow therapeutics (A831 & A688)	ARRET
	Inhibiteur de NS5A	Bristol-Myers quibb (BMS790052)	II
	Inhibiteur non nulceosidique administré par voie orale	Vertex Pharmaceuticals (VCH-759)	II
NS5B	Inhibiteur nucleosidique	Japan tabacco	II
	Inhibiteur nucleosidique	Idenix Pharmaceuticals (Valopicitabine)	ARRET
	L'inhibiteur nucleosidique oral	Roche (R1626)	ARRET
	Inhibiteur non nucleosidique	ViroPharma (HCV 796)	ARRET
	Inhibiteur nucleosidique	Pharmasset (R7128)	II
	Inhibiteur non nucleosidique	Gilead Sciences GS9190	II
Inhibition de l'entrée du VHC	Immunoglobuline polyclonales	HCIg (Civacir)	II
	Anticorps monoclonaux	HCV AB68	ARRET
	Inhibition de l'interaction entre les glycoprotéines virales et les récepteurs cellulaire	Lectine-cyanovirin-N	Pre-clinique
	Inhibition de l'entré du virus dans les cellules hépatiques	iTherX Pharmaceuticals (ITX5061)	II (2009)
E1	Vaccin thérapeutique	innogenetics	II
Voie IMPDH	Inhibiteur	Vertex Pharmaceuticals	III
		Ribapharm	I
		Ribapharm	I
Fibrose du foie	Anti-fibrotique	Intermune Pharmaceuticals	II
		Intermune Pharmaceuticals	III
Apoptose du foie	Inhibiteur des caspases	Idun Pharmaceuticals	II
HCC	Inhibiteur de la b-tubuline	Tularik	III

Ré-infection du VHC	Immunoglobuline anti HCV	Nabi Pharmaceuticals	II
	activité immunosuppresseur	Roche Holdings	Pre-clinique
Stimulation de la réponse immunitaire	Prometteur pour les patients difficiles à traiter	IFN consensus (Alfacon)	IV
	Association intéressante avec le CIFN	IFN gamme 1b (Actimmune)	II
	Dose en continu assez efficace	IFN omega (Intarcia)	II
	Dose bi mensuelle, même effet que IFN-PEG	Albinterferon alfa 2 (Albuferon)	III
	IFN longue durée et administré par voie orale	Interferon oral (Belorofon)	I
	IFN à libération contrôlé	BLX-883 (Locteron)	II
	Inducteur de l'IFN alfa	CPH 10101 (Actilon)	I
	Inducteur de l'IFN alfa	ANA245 (Isatoribine)	ARRET
	DEBIO 025	Utilisé pour les co-infectés	II
	Viramidine	Similaire à la Ribavirine avec moins d'effet secondaire	III
	NIM 811	n.a	I
	SCY 635		I
	DEBIO 10		I
Inhibiteurs de la traduction de l'ARN viral	ISIS 14803	oligonucléotide antisens, problème d'augmentation des ALAT	ARRET
	AVI 4065	oligonucléotide antisens peu efficace	ARRET
	Heptazyme	Ribozymes, problème de toxicité	ARRET
	VGX 410C	Bloque la formation du complexe eIF3/IRES	ARRET
	TT033	ARN interférant, dégradant l'ARN viral	Pré-clinique
Inhibiteurs de la réplication	NM283 (Valopicitabine)	na.	ARRET
	R1626		ARRET
	HCV 796		ARRET
	R7128		II
	GS9190		II
Inhibiteur de l'assemblage et du relargage du virion	PF 03491390	inhibiteurs de caspases	II
	GS9450	inhibiteurs de caspases	II
	CTS1027	Anti inflammatoires	II
	MitoQ	Anti inflammatoires	II
	Nitazoxanide (Alinia)	Bloque la synthèse des protéines virales	II
	Bavituximab (Tarvacin)	Anti phosphotidylsérines	II

Annexe II : Liste des oligonucléotides générés pour le clonage

OLIGONUCLÉOTIDES	SÉQUENCE
CLONAGE CLASSIQUE	
MRR-eIF3b For NdeI ATG	AGTGCGCATATGTCGGTGATTGTAGTGGACAAT
MRR-eIF3b Rev STOP XHOI	TGCTGACTCGAGTTAATCCGTAAAGAGGTTGAC
MRR-eIF3b V2-For	GCTCATATGTCGGTGATTGTAGTGGAC
MRR-eIF3b V2-Rev	AGCCTCGAGATCCGTAAAGAGGTTGACC
NewMRRceIF3bFor	GCTCCATGGTCCGGTGATTGTAGTGGAC
NewMRRceIF3bRev	AGCCTCGAGTCATTAATCCGTAAAGAGGTTGACC
P44 MRR For NdeI ATG	AGTGCGCATATGGCCACCATCCGTGTCACC
P44 MRR Rev STOP XhoI	TGCTGACTCGAGTTACGGCTTGGCCCACTCGAC
p44 ATG NdeI For	AGTGCGCATATGCCTACTGGAGACTTTGAT
p44 XhoI STOP Rev	TGCTGACTCGAGTTAGTTGGTGGACGGCTT
p44 EcoRI ATG For	AGTGCGGAATTCCATGCCTACTGGAGACTTTGAT
p44 full NcoI ATG For	AGTGCGCCATGGGAATGCCTACTGGAGACTTTGAT
Nter eIF3j For	GCTCATATGGCGGGGGACTCGGACTCCTGG
Nter eIF3j Rev	TCATGCTCGAGTCATTTACTTCTGCTTCCTC
eIF3j For	GATCATATGGCGGGGGACTCGGACTCCTGGGA
eIF3j Rev	TCATGCTCGAGTCACATGAAGTCTTCATAGTC
eIF3b full EcoRI ATG For	AGTGCGGAATTCCATGCAGGACGCGGAGAAC
eIF3b full NcoI ATG For	AGTGCGCCATGGGAATGCAGGACGCGGAGAAC
eIF3b full NdeI ATG For	AGTGCGCATATGCAGGACGCGGAGAAC
eIF3b full STOP XhoI Rev	TGCTGACTCGAGTCACTCCTGATTCCCGAG
CLONAGE TECHNIQUE LIC	
LICeIF3aV1For	ATGGGAAGTGAATGCCGGCCTATTTTCAGAGGCCGAAAATGCC
LICeIF3aV1Rev	TTAGAGGGTTAGGGACTCAACGTCGTAAGTGGTCCATCCATCTT
LICeIF3aV2Rev	TTAGAGGGTTAGGGACTCAATCGCCAAGTCTTCTCTCTCTCTACGCC
LICeIF3aV3For	ATGGGAAGTGAATGCTTGAGGAATTGGATCCAGATTTTAT
LICeIF3aV3Rev	TTAGAGGGTTAGGGACTCACTCTTCTCTCGGTTTTGCTCGTTC
LICeIF3aV4For	ATGGGAAGTGAATGCGCGATAGTTCCCTTCTAGAAAGGACTCTC
LICeIF3b D1-134 For	ATGGGAAGTGAATGCGAGCGCCGAGGCCGAAC
LICeIF3b D1-276 For	ATGGGAAGTGAATGCAGAGTGACGAGTGGGATATCCAGAG
LICeIF3b D1-702 For	ATGGGAAGTGAATGCTGCTGTGGCGGCCCCGGCCT
LICeIF3b D133-814	TTAGAGGGTTAGGGACTCAGCCCTCGTTTCTCCCGC
LICeIF3b D276-814 Rev	TTAGAGGGTTAGGGACTCACTGATCGTCATATACTT
LICeIF3b D702-814	TTAGAGGGTTAGGGACTCAGCTGGCAGAAGCGGTCCTTG
LICeIF3b LIC For	ATGGGAAGTGAATGCAGGACGCGGAGAACGTGG
LICeIF3b LIC Rev	TTAGAGGGTTAGGGACTCACTCCTGATTCCCGAGG

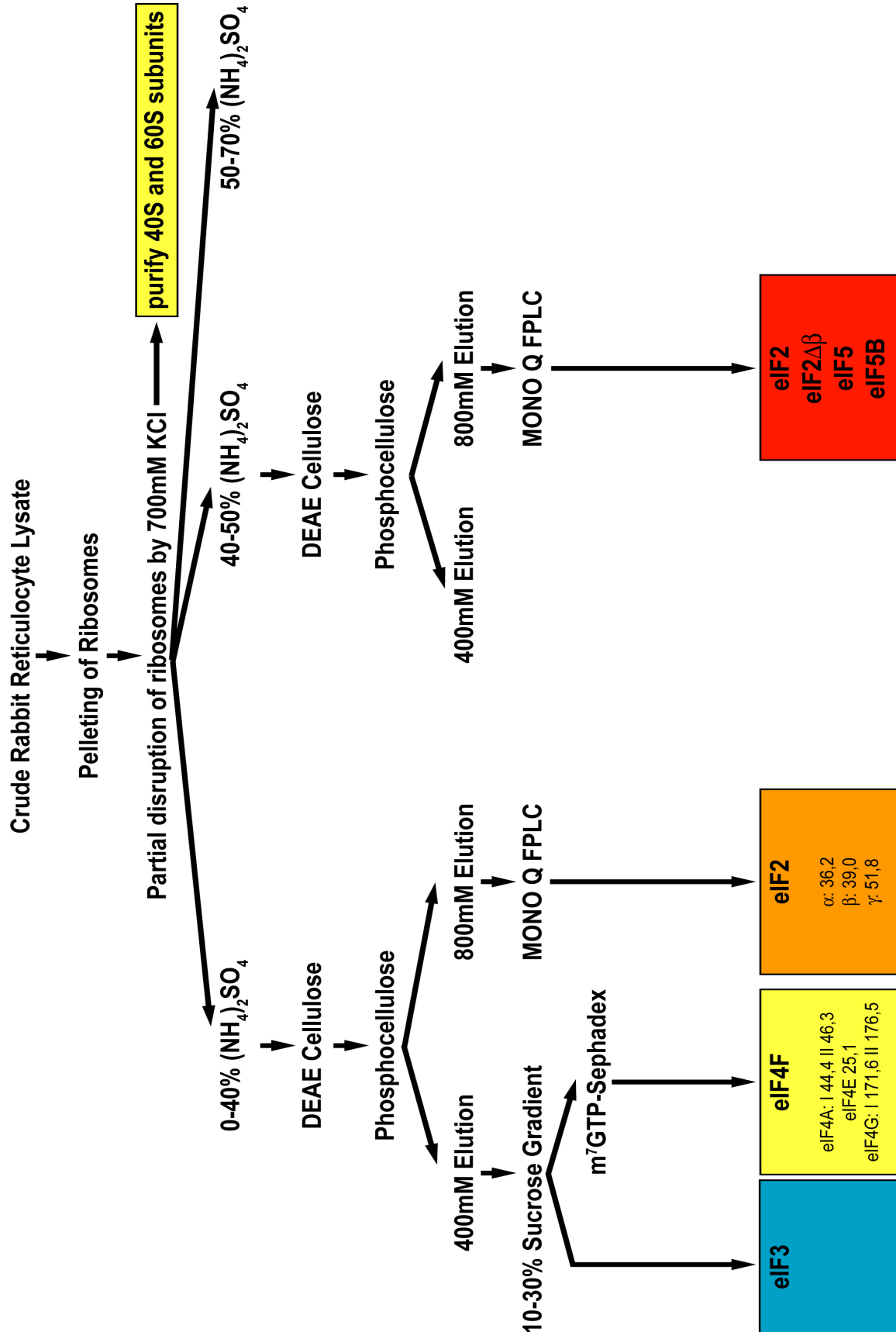
LIceIF3b LIC1 Full For	ATGGGAAGTGAATGCAGGACGCGGAGAACGTGGCGG
LIceIF3b LIC2 Rev	TTAGAGGGTTAGGGACTACTCTGATTCCCGAGG
LIceIF3bV11Rev	TTAGAGGGTTAGGGACTCAGGGCCGCCACAGCAGCTGGCAGAAG
LIceIF3bV11For	ATGGGAAGTGAATGCCGGTGATTGTAGTGGACAATGTCCTCAGGTGG
LIceIF3cV1For	ATGGGAAGTGAATGCCGCGTTTTTCACCACCGTTCCGACAG
LIceIF3cV1Rev	TTAGAGGGTTAGGGACTCAGTAGCCGCTCTGAGACTGCTGCTGGCG
LIceIF3cV2For	ATGGGAAGTGAATGCCCAAGAGAGTTGTCCGCAGTGCCAAGGACAA
LIceIF3cV2Rev	TTAGAGGGTTAGGGACTCAGGGATTGCAAACAGGATATCCATCAG
LIceIF3cV3For	ATGGGAAGTGAATGCCTGCATCTAACTCTGGTGAACGAATGG
LIceIF3cV3Rev	TTAGAGGGTTAGGGACTACTGCTTGTGGTCAAACACCCGTTCC
LIceIF3cV4For	ATGGGAAGTGAATGCCGAAGTGCAGATCACCCATGCTGTTGTAT
LIceIF3dV2For	ATGGGAAGTGAATGCGTAGTGATTGGGAAGTAAAGAGGAAATG
LIceIF3dV2Rev	TTAGAGGGTTAGGGACTCAGCCATCAGGGAGGCTGTAGA
LIceIF3dV1For	ATGGGAAGTGAATGCTGGCAAAGTTCATGACACCCGTGAT
LIceIF3dV1Rev	TTAGAGGGTTAGGGACTCAACGAACTTCAACTGAAGAGTCTCGG
LIceIF3dV0Rev	TTAGAGGGTTAGGGACTCAATCTTCATCAGAGCTGAAGGTGC
LIceIF3eV0For	ATGGGAAGTGAATGCTGGCGGAGTACGACTTGACTACTCGCATC
LIceIF3eV0Rev	TTAGAGGGTTAGGGACTCAGTAGAAGCCAGAATCTTGAGTTGC
LIceIF3eV1Rev	TTAGAGGGTTAGGGACTCACAGTCTTGCAATTTCTAATCAAATTTACA
LIceIF3fV1For	ATGGGAAGTGAATGCTGGCCACACCGCGGTACCAGTAAGTCTCTCC
LIceIF3fV1Rev	TTAGAGGGTTAGGGACTCACAGGTTTACAAGTTTTTCATT
LIceIF3fV2For	ATGGGAAGTGAATGCGCGTGGTCAGGCTGCACCCAGTCATTTGGCCTCC
LIceIF3hV1For	ATGGGAAGTGAATGCCGTCCCGCAAGGAAGTACC GGCTCTA
LIceIF3hV1Rev	TTAGAGGGTTAGGGACTGTTGTTGATTCTTGAAGAGCCTGGGCCATGA
LIceIF3hV2For	ATGGGAAGTGAATGCAAGTGCAGATAGATGGCCTTGTGGTATTAAGA
LIceIF3irev	TTAGAGGGTTAGGGACTCAAGCCTCAAACCTCAAATTCGAAGTACTGTG
LIceIF3iFor	ATGGGAAGTGAATGCTGAAGCCGATCCTACTGCAGGGCCATGAGCG
LIceIF3jFor	ATGGGAAGTGAATGCCGGGGACTCGGACTCCTGGGACGCCGAC
LICNter-elf3jRev	TTAGAGGGTTAGGGACTCATACTTCTGCTTCTCTTTTTTTTCATC
LIceIF3jRev	TTAGAGGGTTAGGGACTCACATGAAGTCTTCATAGTCTTGACATA
LIceIF3EIV1For	ATGGGAAGTGAATGCTGTCTTATCCCGCTGATGATTATGAGTC
LIceIF3EIV1Rev	TTAGAGGGTTAGGGACTCAAGTCTCTGTCCCATCTTCTCAGGGTT
LIceIF3EIV2For	ATGGGAAGTGAATGCCAAAGCAGGACCTTGCTTATGAACGTCAAGTAT
LIceIF3EIV2Rev	TTAGAGGGTTAGGGACTCAATCTTCTCAGGGTTTCGATTAAGCTCCTCA
LIceIF3MV1For	ATGGGAAGTGAATGCTGAGCGTCCCGGCTTCATCGACATCAGT
LIceIF3MV1Rev	TTAGAGGGTTAGGGACTCAGGTATCAGAAAGACTCAAAAGGCT
LIceIF3MV2For	ATGGGAAGTGAATGCATTTAGCTCAAATATTGAAGCCTGTGATGTGT
LIceIF3MV2Rev	TTAGAGGGTTAGGGACTCATCTCTGGGTCTGATCAATTTTGCAGTAG
pBADLICRev	CATTCCACTTCCCATGGCGCCCTGAAAATAAAG
pBADLICFor	TCCCTAACCTCTAAGATCCGAATTCGAGCTCC

pETTEVLIC For	TCCCTAACCCCTCTAAGAGCTCCGTCGACAAGCT
pETTEVLICRev	CATTCCACTTCCCATGTGGCCCTGAAAATACAG
pGEXTEVLICFor	TCCCTAACCCCTCTAAGACTGACGATCTGCCTCG
pGEXTEVLICRev	CATTCCACTTCCCATGACCCGGGAATTCGGGGA
pETM40LICfor	CAT TCC ACT TCC CAT GGC GCC CTG AAA ATA AAG
CLONAGE CLASSIQUE IRES	
Domaine II, III, IV (39-370)	For : CCGGAATTC AATTAATACGACTCACTATAGGGACTCCCCTGTGAGGAACTACTGTCTTCACG Rev : CCCAAGCTTTTCTTTGAGGTTTAGGATTTGTGCTCATG
Domaine II, III (39-330)	For : CCGGAATTC AATTAATACGACTCACTATAGGGACTCCCCTGTGAGGAACTACTGTCTTCACG Rev : CCCAAGCTTTACGAGACCTCCCGGGCACTCGCAAGCAC
Domaine III, IV (134-370)	For : CCGGAATTC AATTAATACGACTCACTATA GGGAGAGCCATAGTGGTCTGCGGAACCGGTGAG Rev : CCCAAGCTTTTCTTTGAGGTTTAGGATTTGTGCTCATG
Domaine III (134-330)	For : CCGGAATTC AATTAATACGACTCACTATAGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAG Rev : CCCAAGCTTTACGAGACCTCCCGGGCACTCGCAAGCAC
Domaine III abcd (134-290)	For : CCGGAATTC AATTAATACGACTCACTATAGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAG Rev : CCCAAGCTTAGTACCACAAGGCCTTTCGCGACCCAACAC
Domaine IIIabc (141-252)	For : CCGGAATTC AATTAATACGACTCACTATAGGGTAGTGGTCTGCGGAACCGGTGAGTACACCG Rev : CCCAAGCTTTAGCAGTCTCGCGGGGCACGCCAAATCT
Domaine III boucle b (187-210)	For : GGAATTC AATTAATACGACTCACTATAGGGCGGGTCTTTCTTGATCAACCCGAAGCTTCGGG Rev : CGAAGCTTCGGGTTGATCCAAGAAAGACCCGCCCTATAGTGTGAGTATTAATTGAATTCGGG
Domaine II, III abcd (39-290)	For : CCGGAATTC AATTAATACGACTCACTATAGGGACTCCCCTGTGAGGAACTACTGTCTTCACG Rev : CCCAAGCTTAGTACCACAAGGCCTTTCGCGACCCAACAC
Domaine II, III abc (39-252)	For : CCGGAATTC AATTAATACGACTCACTATAGGGACTCCCCTGTGAGGAACTACTGTCTTCACG Rev : CCCAAGCTTTAGCAGTCTCGCGGGGCACGCCAAATCT
Domaine II (43-120)	For : CGGGAATTC AATTAATACGACTCACTATAGGGCCTGTGAGGAACTACTGTCTTCAC Rev : ctcaagcttctggaggctgcacaca
CLONAGE GENE SYNTHÉTIQUE IRES	
dIIIdefIVv2	IRES3def4V2a : GGGCCTTAAGATTATGCTGAGTGATATCCCGGAGGGCCCTCTCGGTATCA IRES3def4V2b : CCTCCCGGAGAGCCATAGTGGTTCGCTGCTAGCCGAGTAGTGTGGGCG IRES3def4V2c : CTCATCACAACCCGCGCTTTCGGAAACACCATGACGGCTACCCACGAACG IRES3def4V2D : GATGGGTGCTTGCACTGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAC IRES3def4V2E : ATCTGGCACGTGGTACTGTGTTTAGGATTTGAGTCTTTCTTTTCGAACCG
dIIIdefv2	IRES3def4V2a : GGGCCTTAAGATTATGCTGAGTGATATCCCGGAGGGCCCTCTCGGTATCA IRES3def4V2b : CCTCCCGGAGAGCCATAGTGGTTCGCTGCTAGCCGAGTAGTGTGGGCG IRES3def4V2c : CTCATCACAACCCGCGCTTTCGGAAACACCATGACGGCTACCCACGAACG IRES3def4V2g : GATGGGTGCTTGCACTGCCCCGGGAGGTCTCGTAAAGCTTGCC IRES3def4V2h : CACGGGGCCCTCCAGAGCATTTCGAACCG IRESt7 : CCCGGAATTC AATTAATACGACTCACTATAGGG

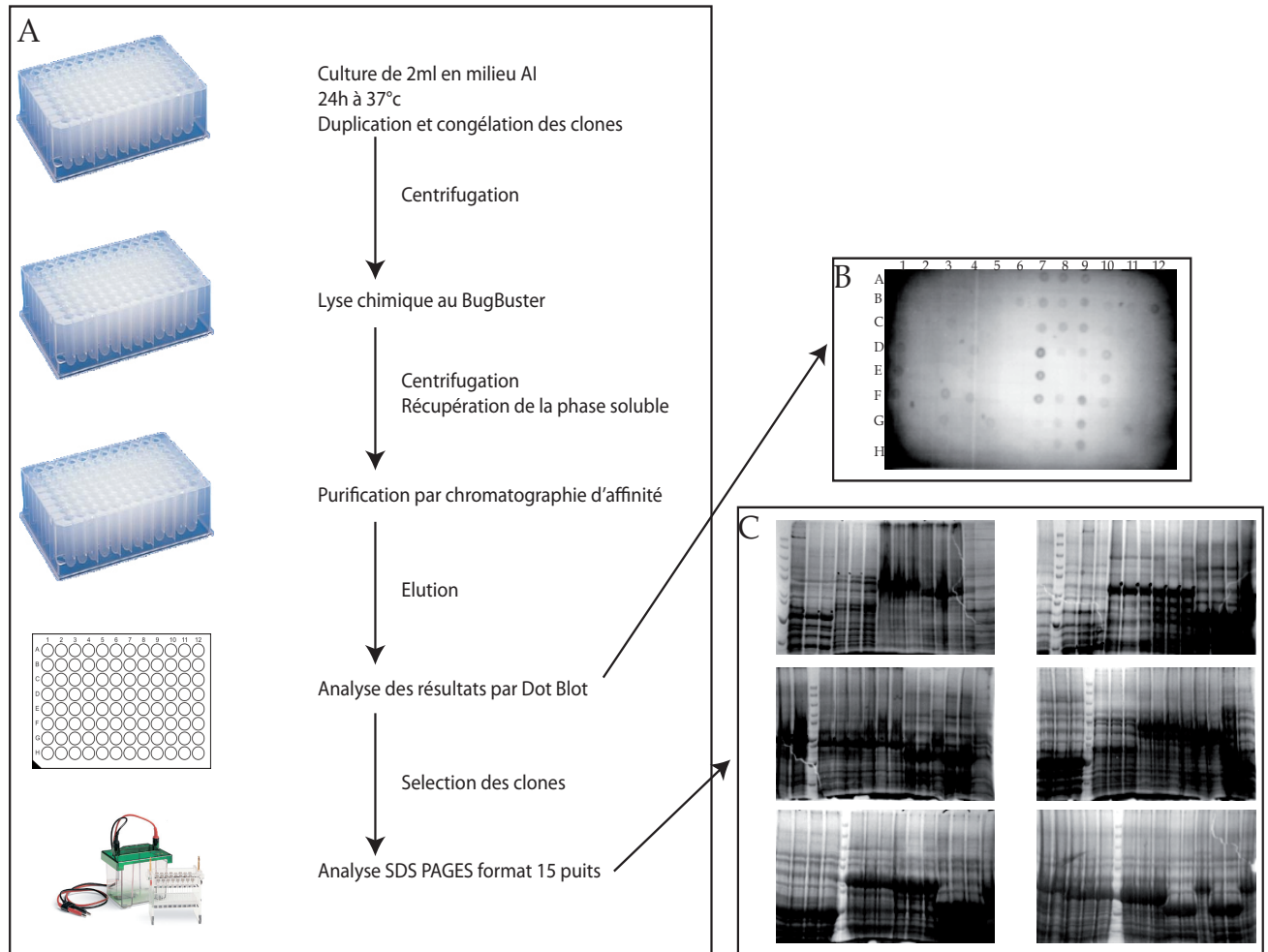
dIIIdefIVv2 -3'	<p>IRES3def4V2a : GGGCCTTAAGATTATGCTGAGTGATATCCCGAGGGCCCTCTCGGTATCA</p> <p>IRES3def4V2b : CCTCCCGGGAGAGCCATAGTGTTTCGCTGCTAGCCGAGTAGTGTGGGGC</p> <p>IRES3def4V2c : CTCATCACAAACCCGCGCTTTCCGGAACACCATGACGGCTACCCACGAACG</p> <p>IRES3def4V2d : GATGGGTGCTTGACAGTGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAC</p> <p>IRES3def4V2f : ATCTGGCACGTGGTACTGTGTTTAGGATTCTGAACCG</p> <p>IRES_{t7} : CCCGGAATTCTAATACGACTCACTATAGGG</p>
dIIIefIVv2	<p>IRES3def4V2i : GGGCCTTAAGATTATGCTGAGTGATATCCCGAGGGCCCTCTCGGTAAGC</p> <p>IRES3def4V2j : CCTCCCGGGAGAGCCATTCGTGGTACTGCCGATGGGTGCTTGCG</p> <p>IRES3def4V2k : ACCATGACGGCTACCCACGAACGCTCACGGGGC</p> <p>IRES3def4V2l : AGTGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAC</p> <p>IRES3def4V2e : ATCTGGCACGTGGTACTGTGTTTAGGATTTGAGTTTCTTTTCGAACCG</p> <p>IRES_{t7} : CCCGGAATTCTAATACGACTCACTATAGGG</p>
dIIIefv2	<p>IRES3def4V2i : GGGCCTTAAGATTATGCTGAGTGATATCCCGAGGGCCCTCTCGGTAAGC</p> <p>IRES3def4V2j : CCTCCCGGGAGAGCCATTCGTGGTACTGCCGATGGGTGCTTGCG</p> <p>IRES3def4V2k : ACCATGACGGCTACCCACGAACGCTCACGGGGC</p> <p>IRES3def4V2l : TTGCGAGTGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAC</p> <p>IRES3def4V2m : ATCTGGCACGTGGTACTGTGTTTAGGATTCTGAACCG</p> <p>IRES_{t7} : CCCGGAATTCTAATACGACTCACTATAGGG</p>
Domaine IIIef, IV (250-370)	<p>IRES₁ : TACCCGGAATTC AATTAATACGACTCACTATAGGGCTCCCGGATGGGT</p> <p>IRES₂ : CTCCCGGGCACTCGCAAGCACCCATCAGGCATCCCGGGAGGCCCTAT</p> <p>IRES₃ : TGCTTGCAGTGCCCCGGGAGGTCTCGTAGACCGTGCATCATGAGCACA</p> <p>IRES₄ : CGGAAGCTTAATTTCTTTGAGGTTTAGGATTTGTGCTCATGATGTCACG</p>
Domaine III boucle b « long » 181 (181-218)	<p>IRES_{B1} : CCGGAATTC AATTAATACGACTCACTATAGGGGATGA</p> <p>IRES_{B2} : ATCCAAGAAAGGACCCGGTCATCCCCTATAGTGAGTCGTAT</p> <p>IRES_{B3} : CCGGGTCTTTCTTGATCAACCCGCTCAATGCAAGCTTCC</p> <p>IRES_{B4} : GGAAGCTTG CATTGAGCG</p>
Domaine III boucle ab 172 (172-227)	<p>IRES_{A1} : CCGGAATTC AATTAATACGACTCACTATAGGGAATTGC</p> <p>IRES_{A2} : AAAGGACCCGGTCATCCTGGCAATTCCTATAGTGAGTCGT</p> <p>IRES_{A3} : AGGATGACCGGGTCTTTCTTGATCAACCCGCTCAATGCC</p> <p>IRES_{A4} : GGAAGCTTAATCTCCAGGCATTGAGCGGGTTGA</p>

Annexe III : Protocole de purification du complexe eIF3 à partir de réticulocytes de lapin.

Source Jan Medenbach, EMBL heidelberg.



Annexe IV : Protocole expérimentale utilisé pour le test d'expression à haut débit des différentes constructions des sous-unités du complexe eIF3.



D

Nom	Version	Construction	Localisation	taille	Masse	Pi	lg/l =	Seq	Clonage			Expression			Purification soluble		
									pGEXTEV	pETTEV	MBP	pGEXTEV	pETTEV	MBP	pGEXTEV	pETTEV	MBP
eIF3a	V1	1 to 1382	MPAYF-TTVRR	1382	166569.3	6.38	1.106	ok		1A01	1C10		1A01	1C10		1A01	1C10
	V2	1 to 880	MPAYF-RRLGD	880	105744.8	7.11	0.754	ok			2E2			2E2			
	V3	661 to 840	LEEL-REEE	180	23	5,46	0,57	OK	2D08	1B04	1E10	2D08	1B04	1E10	2D08	1B4	1E10
	V4	880 to 13880	DSSL-TTVRR	522	63,4	5,5	1,64	ok	2E08	1E01	2G2	2E08	1E01	2G2	2E08	1E1	2G2
eIF3c	V1	1 to 913	MPRF-QTAY	913	105343.8	5.48	0.804										
	V2	45 to 416	PKRV-FANP	371	42432.4	5.12	0.8	ok		1G02	1F11		1G02	1F11		1G2	1F11
	V3	442-878	RCIL-DHKQ	436	50977.4	6.11	0.801										
	V4	320-878	RGTEI-DHKQ	558													
eIF3d	V0	1 to 547	LAKF-SDED	532	61934.9	8.74	1.403	ok	2B10	1H05	1A12	2B10	1H05	1A12	2B10	1H05	1A12
	V1	1 to 164	LAKF-VEVR	164	19502.9	10.10	1.381	ok	2D10		1D12	2D10		1D12	2D10		1D12
	V2	164 to 525	RSDW-LPDG	360	41512.0	5.81	1.445	ok			1G12			1G12			1G12
eIF3e	V0	1 to 445	LAEY-SGFY	445	52220.8	5.71	1.210	ok	2G10	1C03	1H12	2G10	1C03	1H12	2G10	1C03	1H12
	V2	1 to 380	LAEY-NARL	380	44935.6	5.42	1.217	ok		1B07	2G03		1B07	2G03		1B07	2G03
eIF3f	V1	1 to 357	MATP-LVNL	357	37.7	5.24	0.58	ok	2C11		2B01	2C11		2B01	2C11		2B01
	V2	90 to 360	RVVR-LVNL	268	29.7	5.35	0.76	ok	2F11	1A06	2E01	2F11	1A06	2E01	2F11	1A06	2E01
eIF3h	V1	1 to 352	MASR-EYNN	352	39.9	6,1	0,798	ok	2G11	1F08		2G11	1F08		2G11	1F08	
	V2	38 to 352	QVQI-EYNN	314	36	5,35	0,76	ok	2D12	1D03	2A02	2D12	1D03	2A02	2D12	1D3	2A2
eIF3i	V1	full	MKPI-EFEA	320	36	5,38	1,2	ok	2D07	1H09	2C05	2D07	1H09	2C05	2D07	1H9	2C5
eIF3E1	V1	full	MSYP-GQRP	325	36,5	5,38	1,2	ok	2H06	1c09		2H06	1c09		2H06	1C9	
	V2	30 to 565	PKQD-GQRP														
eIF3M	V1	full	MSVP-LSDT	531	63	6,76	1,5	ok	2A07	1E09	2F04	2A07	1E09	2F04	2A7	1E9	2F04
	V2	41 to 330	DLAQ-QTQR	290	33	5,28	0,7	ok			2A05			2A05		2A05	

Protocole expérimentale utilisé pour le test d'expression à haut débit des différentes constructions des sous-unités du complexe eIF3.

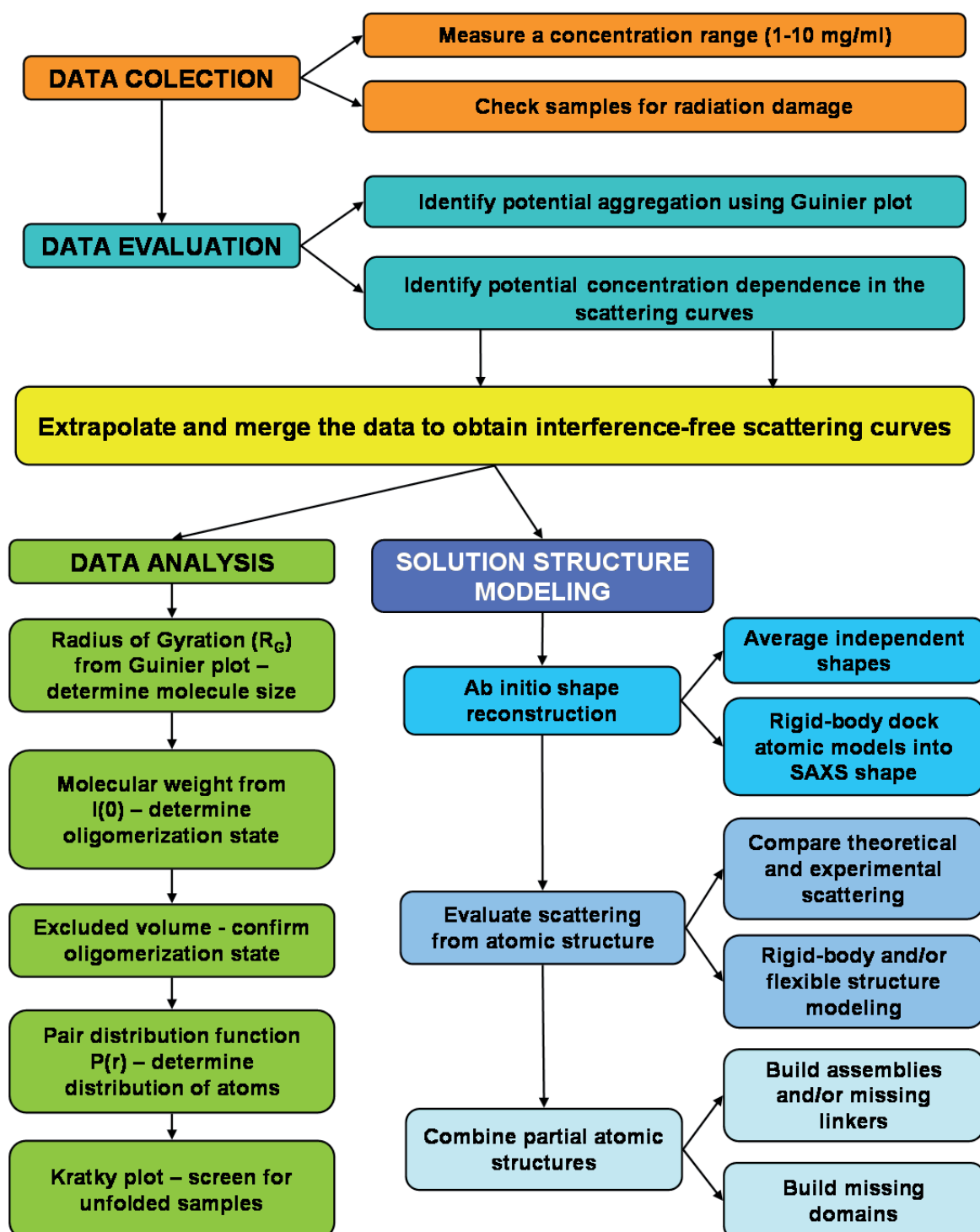
A : Culture, induction et lyse en format 96 puits.

B : Détection de l'expression et de la solubilité par DOT BLOT.

C : Gel SDS PAGE de 3 clones de chaque constructions soluble.

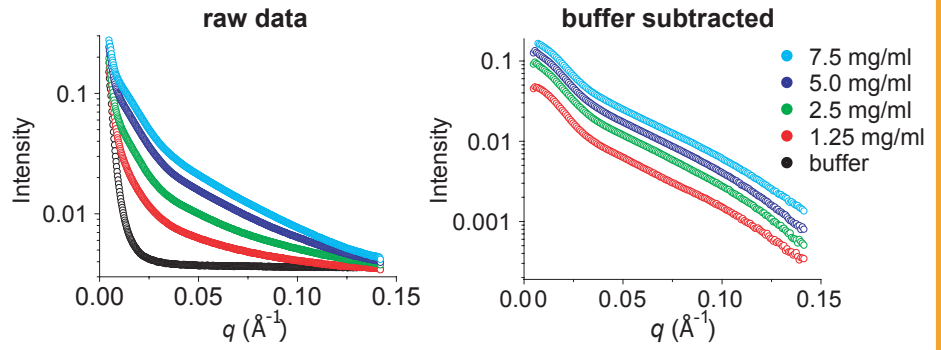
D : Tableau récapitulatif des clonages et test d'expression et de solubilité pour chaque constructions testées

Annexe V : Protocole expérimentale présentant les différentes phases de la technique de SAXS (Svergun).

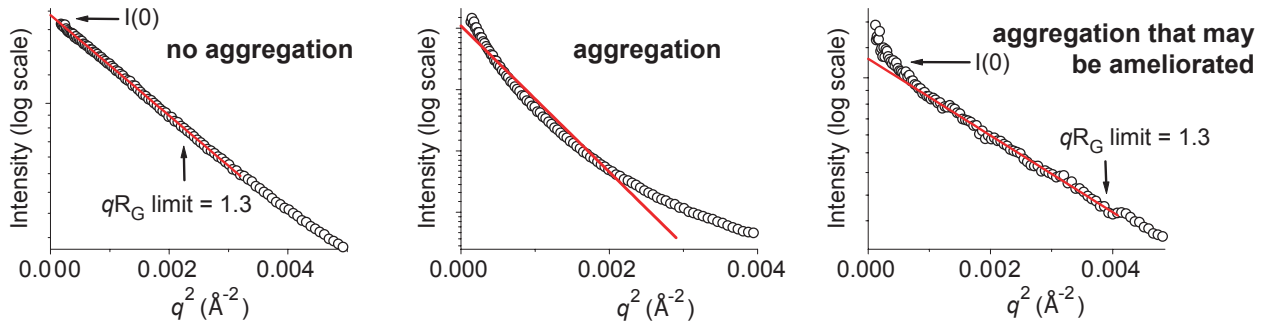


DATA COLLECTION

Measure four different concentrations in the range 1-10 mg/ml. X-ray sensitivity can be detected by changes in scattering observed in repeated exposures and typically lead to increases in R_G and $I(0)$. X-ray damage can be minimized by adding 5-10% glycerol.

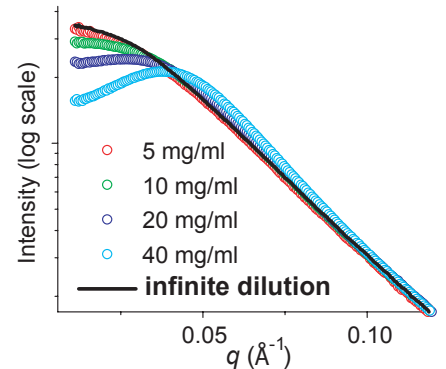


Calculate R_G and $I(0)$ from the Guinier plot. A non-linear dependence of $\log(I(q))$ vs. q^2 indicates the presence of aggregation. **Scattering from aggregated samples strongly influences the entire data set and no further processing should be performed!** Aggregation in the sample can be reduced or eliminated by varying buffer conditions, centrifugation, and filtration.

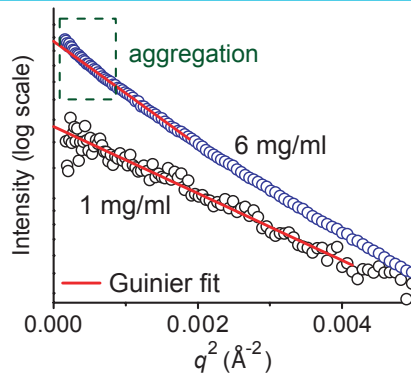


DATA EVALUATION

Evaluate the effects of interparticle interference $F(q)$ with the scattering of individual particles $S(q)$. Superposition of scaled scattering curves from multiple concentrations can reveal concentration-dependent scattering. Decreases in intensity at very small q with increasing protein concentration indicates interparticle repulsion. Repulsion can be dealt with by dilution, increasing ionic strength, and/or extrapolating the data to infinite dilution. Scaling the data to extrapolated scattering is only required in the range where the interference is visible, typically $q < 0.1 \text{ \AA}^{-1}$.



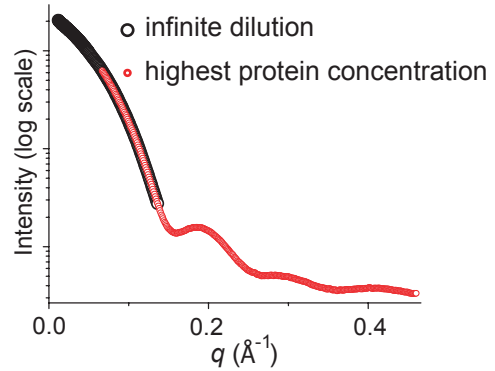
Watch for increasing $I(0)/c$ and R_G with increasing concentration. These can be due to changes in multimerization state or increasing amounts of aggregation in the sample.



sample	R_G	$I(0)/c$
6 mg/ml	30 Å	94
1 mg/ml	22 Å	40

IDEAL SCATTERING

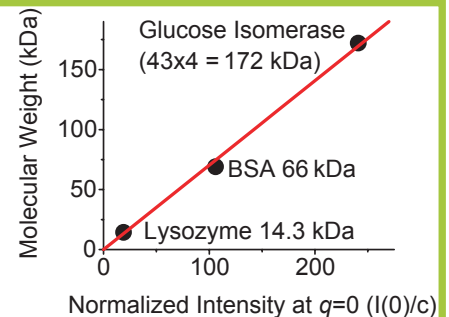
To obtain an ideal scattering curve for the entire q range, the scattering profile must be extrapolated to infinite dilutions at low resolution ($q < 0.1 \text{ \AA}^{-1}$) and merged with the scattering profile for larger angles. Accurate large angle data can be obtained by measuring higher concentrations, using longer exposure times, and/or decreasing the sample-to-detector distance.



If SAXS data are collected from monodisperse samples without interparticle interference, then reconstruction of the solution structure can proceed.

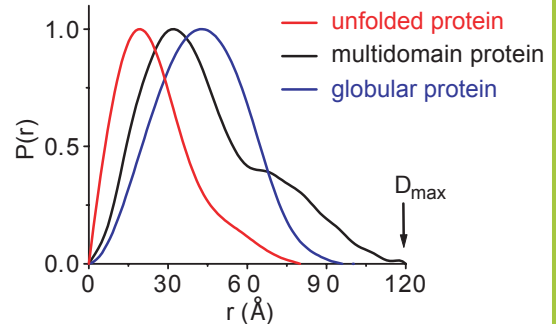
R_G , a measure of the overall size of the molecule, can be determined from the Guinier plot.

Molecular weight from $I(0)$ requires a calibration curve and can be used to determine **oligomerization state**. This can be validated using calculation of the **excluded volume**.

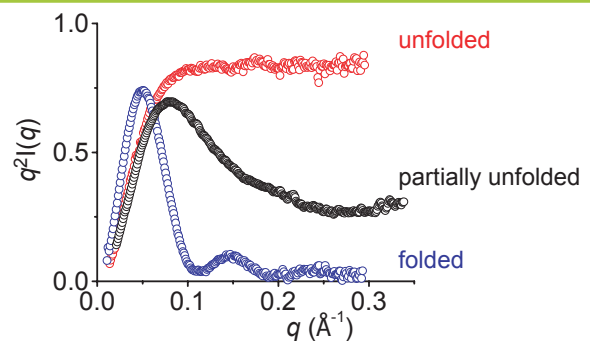


DATA ANALYSIS

Globular macromolecules have a $P(r)$ function with a **single peak**, while elongated macromolecules have a longer tail at large r and can have multiple peaks. The maximum length in the particle, D_{\max} , is the position where the $P(r)$ function returns to zero at large values of r . **Disagreements for values of R_G and $I(0)$** calculated from the $P(r)$ function and from the Guinier plot can indicate small amounts of aggregation that primarily affect the low resolution data and the accuracy of the Guinier plot.



The Kratky plot identifies unfolded samples. Globular macromolecules follow Porod's law and have bell-shaped curves. Extended molecules, such as unfolded peptides, lack this peak and have a plateau or are slightly increasing in the larger q range.



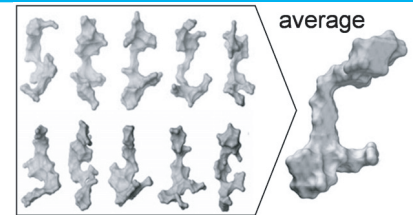
SOLUTION STRUCTURE MODELING

No atomic structure: *ab initio* shape reconstruction

Determine the low resolution envelope with *ab initio* calculations. When the number of residues are known at the data extends to reasonably large values of q ($q_{max} > 0.3 \text{ \AA}^{-1}$), GASBOR is recommended as it provides some advantages for anisotropic shapes.

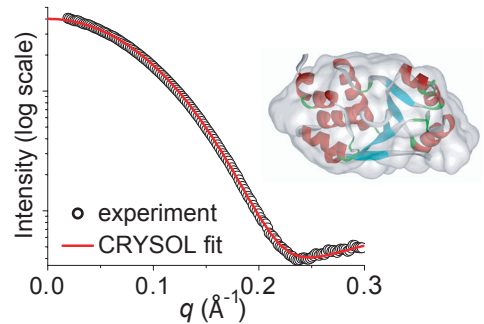


Average *ab initio* envelopes from multiple runs. To avoid problems due to over fitting and to ensure a reproducible solution, calculate multiple individual envelopes and use the averaged solution.

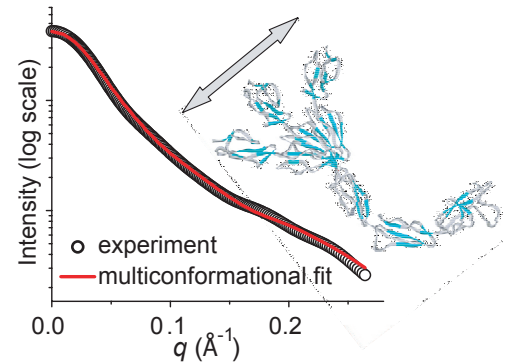


Available atomic structure: evaluate theoretical scattering reconstruction

Compare experimental SAXS scattering to SAXS scattering calculated from atomic structures. Perfect matches indicate that the solution and crystal structures are essentially identical.



Theoretical and experimental scattering can disagree. a) Investigate fitting with alternate assembly states. If symmetry is suspected, try GLOBSYMM. b) Investigate fitting with mixtures of different assembly states. c) Investigate the possibility of flexibility between domains and missing loops that may need to be modeled in a multi-conformational way.

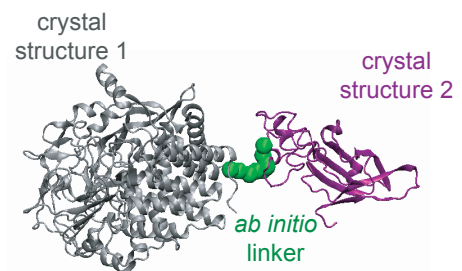


Partial atomic structure: rigid body modelling + *ab initio* reconstruction

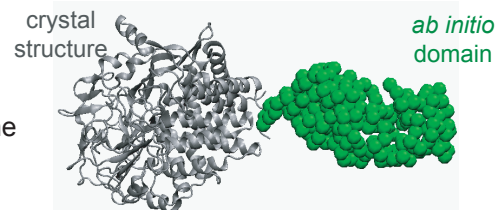
Individual subunits are known but not the overall assembly. Start with rigid body modeling of the protein complex. Quaternary structure modeling can be performed in a fully automated way using SASREF.



Rigid body modeling does not identify candidate assemblies with appropriate scattering and/or linkers between domains are long. Investigate building multidomain models with BUNCH using individual domain structures and potentially combining multiple scattering curves when SAXS data from deletion constructs are available.



Structure of one domain is known, but others are missing. Construct missing domains using CREDO in combination with the known fragment.



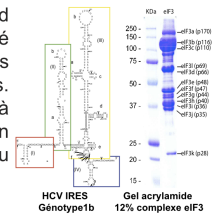
SUPPLEMENTS

La sous unité b du facteur d'initiation eIF3 interagit avec l'IRES du VHC par son motif de reconnaissance à l'ARN (MRR)

Julien Pérard, Rodolfo Rasia, Jan Medenbach, Isabel Ayala, MaríaRosa Quintero Bernabeu, Jérôme Boisbouvier, Emmanuel Drouet & Florence Baudin

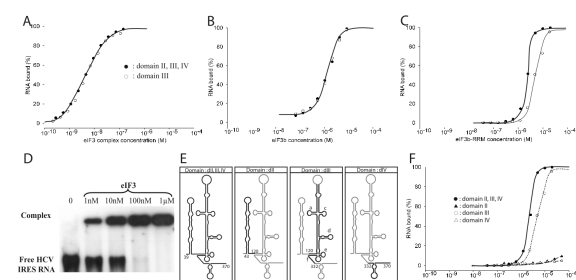
Introduction

L'IRES (Internal Ribosome Entry Site) du VHC est agencé en une structure secondaire complexe hautement conservée qui comprend quatre domaines distincts (I-IV). Cette organisation permet la liaison directe du facteur d'initiation eucaryote eIF3 et de la sous-unité 40S du ribosome au niveau de l'ARN. Cette stratégie permet au virus de court-circuiter le recrutement des autres facteurs eucaryotes au profit de la traduction des ARN viraux. Cependant, les bases moléculaires de l'assemblage de ce complexe restent énigmatiques. Une analyse bio-informatique des différentes sous unités du complexe eIF3 a révélé la présence de deux motifs de reconnaissance à l'ARN (MRR) : l'un au niveau de la sous-unité eIF3b (Aa :185-268) et l'autre sur eIF3g (Aa : 239-317). Nous proposons ici l'association de plusieurs techniques (Filter Binding Assays : FBA, Spectroscopie RMN, Small Angle X-ray Scattering : SAXS) pour l'étude fine du complexe {eIF3b-MRR / IRES VHC}.



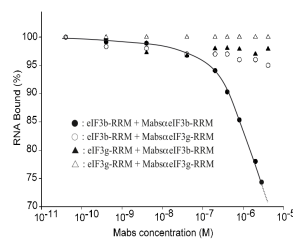
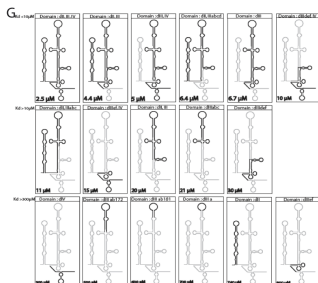
Résultats

Dissection moléculaire de l'interaction eIF3/IRES: Afin de déterminer le rôle de ces motifs, le complexe eIF3, la protéine eIF3b et les différents domaines MRR ont été étudiés par la méthode FBA, en interaction avec différentes constructions d'IRES VHC (cf. fig G). Nous avons pu caractériser les constantes de dissociation entre : eIF3/IRES (Kd 5 nM) (fig A), eIF3b/IRES (Kd 2 μM) (fig B) et eIF3b-MRR (Kd 2 μM) (fig C). Ces résultats confirment l'implication du domaine MRR de eIF3b dans l'interaction eIF3-IRES. De plus, ces résultats complètent les données publiées par d'autres équipes et montrent que eIF3b-MRR se lie directement au domaine III de IRES via son domaine N-terminale (fig E & F).

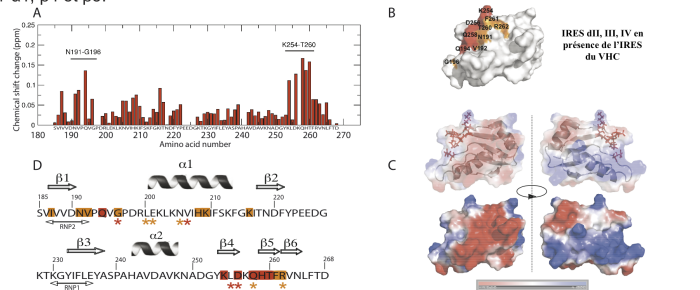


(A) : FBA eIF3/HCV IRES dII, III, IV. Kd = 5nM
(B) : FBA eIF3b/HCV IRES dII, III, IV. Kd = 2μM
(C) : FBA eIF3b-MRR/HCV IRES dII, III, IV. Kd = 2μM
(D) : EMSA eIF3/IRES dII, III, IV. Kd = 10μM
(E) : dissection moléculaire de l'IRES, domaine dII, dIII, dIV
(F) : FBA eIF3b-MRR / domaine dII, dIII, dIV
(G) : Tableau récapitulatif des différentes constructions testées en FBA

Inhibition de la formation du complexe eIF3b-MRR/IRES : Nous avons réussi à inhiber spécifiquement la formation du complexe eIF3b-MRR / IRES HCV grâce à un anticorps monoclonal IgG1 dirigé contre le domaine eIF3b-MRR. (Utilisation d'un rapport 1:1 en concentration molaire pour la réalisation des tests d'inhibition).

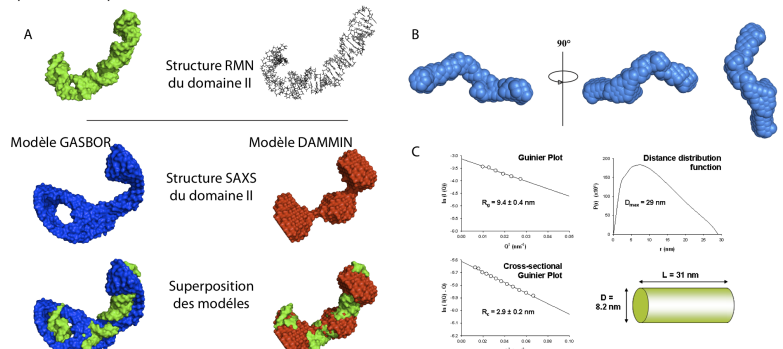


Identification des acides aminés impliqués dans l'interaction eIF3b-MRR/IRES: Après avoir validé le rôle de ce domaine dans l'interaction avec l'IRES du VHC, nous avons utilisé la spectroscopie RMN afin d'identifier les acides aminés impliqués dans la reconnaissance de l'ARN viral. Nous avons pu identifier neuf résidus impliqués dans cette interaction qui sont situés au niveau des structures : β1-α1, β4 et β5.



(A) : Déplacement chimique mesuré lors de la formation du complexe eIF3b-MRR/IRES dII, III, IV. (B) : Enveloppe du domaine eIF3b-MRR en complexe avec l'IRES. (C) : Surface électrostatique du domaine eIF3b-MRR. (D) : schématisation des acides aminés impliqués dans les différentes interactions. Code couleur : rouge = zone d'interaction (sup à 0,1), orange = zone d'interaction (comprise entre 0,05 et 0,1).

Caractérisation de la dimension moléculaire de l'IRES libre en solution par SAXS: la technique SAXS (Small Angle X-Rays Scattering), nous a permis de caractériser pour la première fois la dimension moléculaire de l'IRES libre en solution. Les données de SAXS seront utilisées afin de générer un modèle de l'IRES libre ou en complexe à une résolution quasi atomique. La résolution structurale de ce complexe Ribonucléoprotéique impliqué dans la traduction du VHC aidera à mieux comprendre les phases initiales de la traduction initiale du VHC.



(A) : Etude structurale par SAXS du domaine II de l'IRES du VHC. Utilisation des programmes GASBOR et DAMMIN pour la génération de modèles expérimentaux qui sont superposés à la structure RMN du domaine II.
(B) : Etude structurale par SAXS du domaine II, III, IV de l'IRES du VHC. Utilisation du programme DAMMIN pour la génération de modèle expérimentaux. Données en cours de traitement.
(C) : Données structurales des modèles obtenus pour l'IRES dII, dIII, dIV par SAXS. Données en cours de traitement.

Conclusion & perspectives

L'étude structurale et fonctionnelle de ce complexe impliqué dans l'initiation de la traduction du VHC nous aidera à mieux comprendre les phases initiales du cycle viral. Les résultats que nous avons obtenu ici complètent les résultats publiés par d'autres équipes et serviront de base pour le « drug design » de molécules inhibant spécifiquement la traduction du VHC. C'est dans ce but que nous avons réussi à inhiber la formation de ce complexe à l'aide d'un anticorps dirigé contre le domaine eIF3b-MRR. Ce résultat est très encourageant dans le cadre du « drug discovery ». De plus les résultats obtenus en SAXS vont nous permettre de réaliser un modèle à haute résolution de l'IRES libre ou complexé avec différentes protéines (eIF3, eIF3b ou eIF3b-MRR) pour étudier finement les zones d'interactions mises en jeu lors de la formation du complexe eIF3/IRESHCV in vitro.

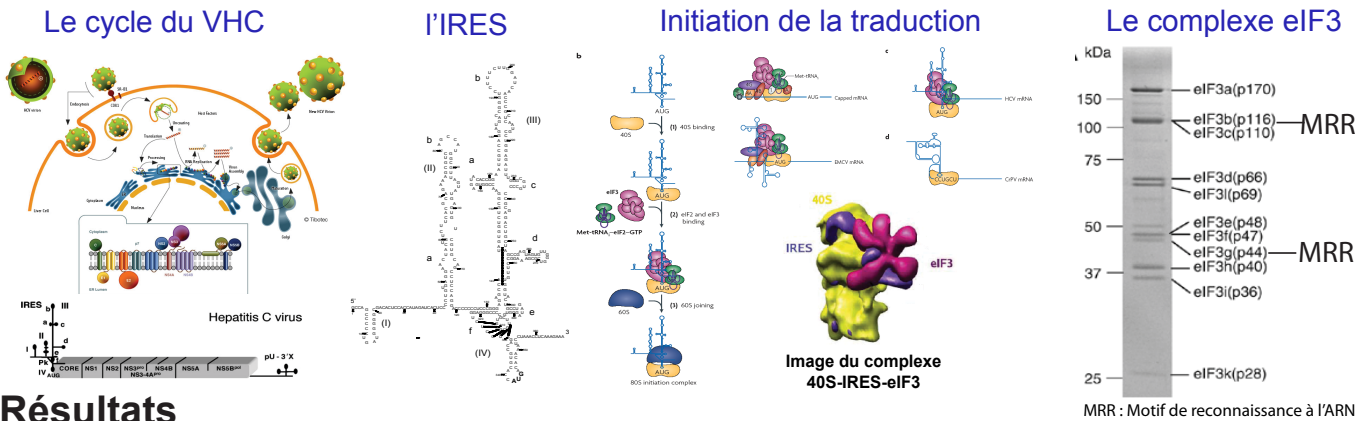
Human initiation factor eIF3 subunit b interacts with HCV IRES RNA through its N-terminal RNA recognition motif. FEBS Letters 583 (2009) 70–74.
Julien Pérard, Rodolfo Rasia, Jan Medenbach, Isabel Ayala, Jérôme Boisbouvier, Emmanuel Drouet, Florence Baudin.

Etude moléculaire de l'initiation de la traduction du Virus de l'Hépatite C en vue d'une nouvelle approche thérapeutique.

Julien Perard, Florence Baudin & Emmanuel Drouet

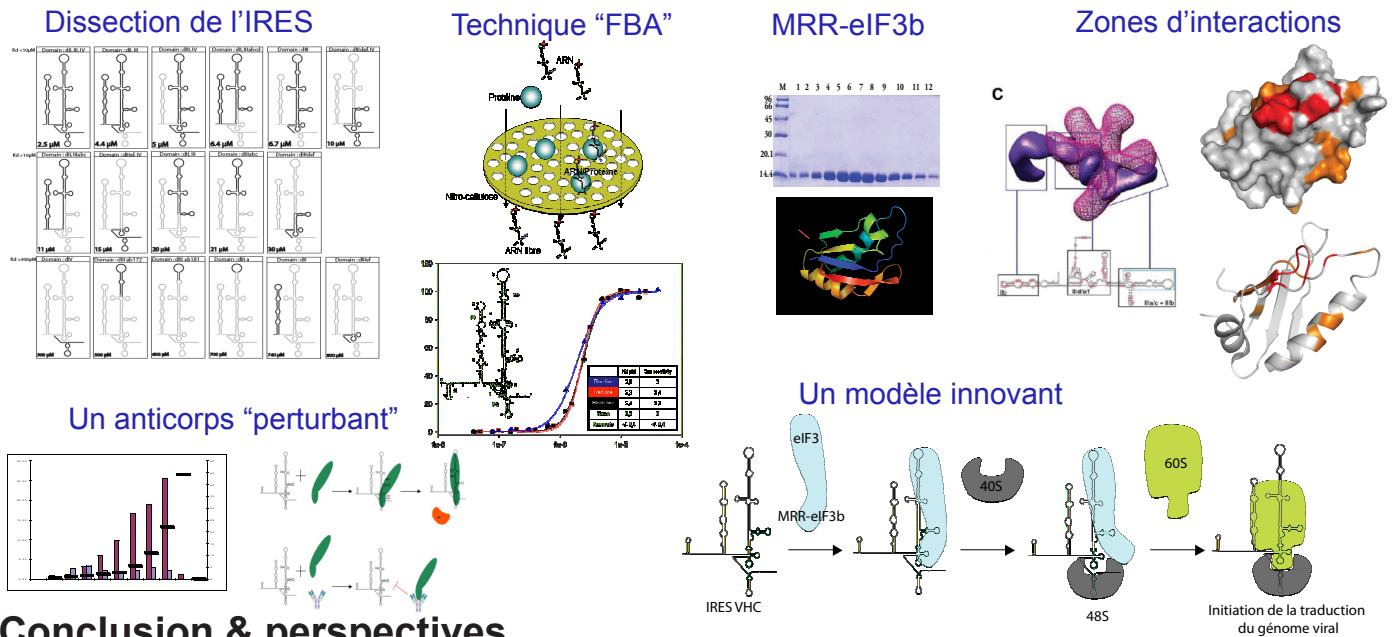
Introduction

Le VHC est une cause majeure d'affection hépatique chronique, notamment la cirrhose et le cancer du foie. On estime que 170 millions de personnes dans le monde sont des porteurs chroniques du VHC et que 3 à 4 millions de personnes sont infectées chaque année. Une caractéristique de ce virus tient à la grande capacité de son génome à muter rendant plus complexe la recherche d'anti-viraux efficaces et susceptibles de prévenir la survenue d'un cancer. La région 5' du génome viral possède un signal d'entrée interne du ribosome ou « IRES » et permet à l'ARN viral d'être traduit en initiant la traduction par le recrutement des sous unités : 40S du ribosome et le facteur d'initiation de la traduction 3 (eIF3). Nos recherches se focalisent sur la formation du complexe entre eIF3 et la région 5' non codante du génome. L'étude biochimique de la formation de ce complexe est une voie très innovante dans la recherche de nouvelles cibles thérapeutiques. Nos objectifs sont de fournir des données biochimiques et fonctionnelles supportant le rôle direct de la région génomique impliqué dans l'interaction avec des protéines cellulaires.



Résultats

Les résultats obtenus en 2006-2007 nous ont permis de mettre en évidence l'existence de complexes stables dont l'affinité et la structure peuvent être identifiées avec précision. À ce jour, nous avons pu identifier une région de 100 nucléotides liant un domaine peptidique du complexe eIF3b. De plus, un anticorps monoclonal dirigé contre un épitope peptidique de ce domaine a permis de réduire l'interaction entre l'ARN viral et la protéine cellulaire, constituant une preuve de concept pour aborder l'inhibition de cette interaction.



Conclusion & perspectives

À l'heure actuelle le traitement anti-VHC est basé sur l'administration de traitement anti-viraux non spécifiques d'une infection au VHC, l'Interféron alpha et à la Ribavirine qui possèdent de nombreux effets secondaires. Les molécules ciblant des complexes formés entre, un ARN viral et une protéine cellulaire eucaryote, sont considérées comme une nouvelle approche, prometteuse et originale. L'étude structurale et fonctionnelle de ce complexe impliqué dans la traduction du VHC aidera à mieux comprendre les phases initiales du cycle viral. Les résultats que nous obtiendrons compléteront les résultats publiés par d'autres équipes et pourront servir de base pour le « drug design » de molécules inhibant spécifiquement la traduction du VHC.

La sous unité b du facteur d'initiation eIF3 interagit avec l'IRES du VHC par son motif de reconnaissance à l'ARN (MRR)

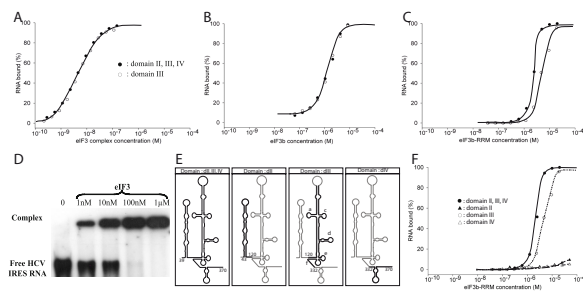
Julien Pérard, Rodolfo Rasia, Jan Medenbach, Isabel Ayala, MaríaRosa Quintero Bernabeu, Jérôme Boisbouvier, Emmanuel Drouet & Florence Baudin

Introduction

L'IRES (Internal Ribosome Entry Site) du VHC est agencé en une structure secondaire complexe hautement conservée qui comprend quatre domaines distincts (I-IV). Cette organisation permet la liaison directe du facteur d'initiation eucaryote eIF3 et de la sous-unité 40S du ribosome au niveau de l'ARN. Cette stratégie permet au virus de court-circuiter le recrutement des autres facteurs eucaryotes au profit de la traduction des ARN viraux. Cependant, les bases moléculaires de l'assemblage de ce complexe restent énigmatiques. Une analyse bio-informatique des différentes sous-unités du complexe eIF3 a révélé la présence de deux motifs de reconnaissance à l'ARN (MRR) : l'un au niveau de la sous-unité eIF3b (Aa : 185-268) et l'autre sur eIF3g (Aa : 239-317). Nous proposons ici l'association de plusieurs techniques (Filter Binding Assays : FBA, Spectroscopie RMN, Small Angle X-ray Scattering : SAXS) pour l'étude fine du complexe {eIF3b-MRR / IRES VHC}.

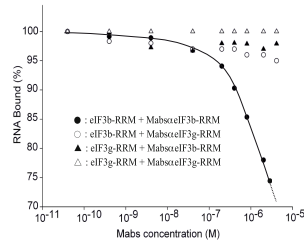
Résultats

Dissection moléculaire de l'interaction eIF3/IRES: Afin de déterminer le rôle de ces motifs, le complexe eIF3, la protéine eIF3b et les différents domaines MRR ont été étudiés par la méthode FBA, en interaction avec différentes constructions d'IRES VHC (cf. fig G). Nous avons pu caractériser les constantes de dissociation entre : eIF3/IRES (Kd 5 nM) (fig A), eIF3b/IRES (Kd 2 µM) (fig B) et eIF3b-MRR (Kd 2 µM) (fig C). Ces résultats confirment l'implication du domaine MRR de eIF3b dans l'interaction eIF3-IRES. De plus, ces résultats complètent les données publiées par d'autres équipes et montrent que eIF3b-MRR se lie directement au domaine III de IRES via son domaine N-terminale (fig E & F).

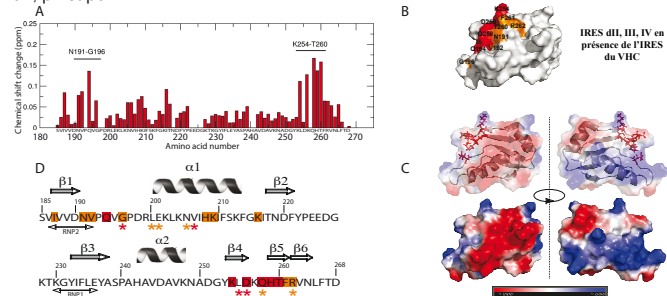


(A) : FBA eIF3/HCV IRES dII, dIII, dIV. Kd = 5 nM
(B) : FBA eIF3b/HCV IRES dII, dIII, dIV. Kd = 2 µM
(C) : FBA eIF3b-MRR/HCV IRES dII, dIII, dIV. Kd = 2 µM
(D) : EMSA eIF3/IRES dII, dIII, dIV. Kd = 10 µM
(E) : dissection moléculaire de l'IRES, domaine dII, dIII, dIV; dII; dIII; dIV
(F) : FBA eIF3b-MRR/domaine dII, dIII, dIV; dII; dIII; dIV
(G) : Tableau récapitulatif des différentes constructions testées en FBA

Inhibition de la formation du complexe eIF3b-MRR/IRES : Nous avons réussi à inhiber spécifiquement la formation du complexe eIF3b-MRR / IRES HCV grâce à un anticorps monoclonal IgG1 dirigé contre le domaine eIF3b-MRR. (Utilisation d'un rapport 1:1 en concentration molaire pour la réalisation des tests d'inhibition).

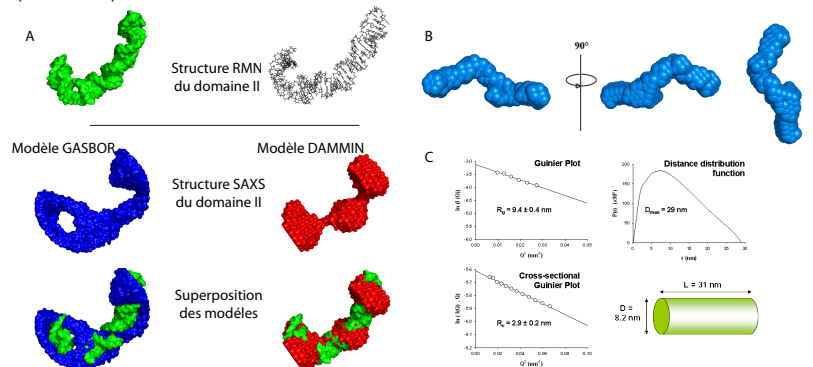


Identification des acides aminés impliqués dans l'interaction eIF3b-MRR/IRES: Après avoir validé le rôle de ce domaine dans l'interaction avec l'IRES du VHC, nous avons utilisé la spectroscopie RMN afin d'identifier les acides aminés impliqués dans la reconnaissance de l'ARN viral. Nous avons pu identifier neuf résidus impliqués dans cette interaction qui sont situés au niveau des structures : β 1- α 1, β 4 et β 5.



(A) : Déplacement chimique mesuré lors de la formation du complexe eIF3b-MRR/IRES dII, dIII, dIV. (B) : Enveloppe du domaine eIF3b-MRR en complexe avec l'IRES. (C) : Surface électrostatique du domaine eIF3b-MRR. (D) : schématisation des acides aminés impliqués dans les différentes interactions. Code couleur : rouge = zone d'interaction (sup à 0,1), orange = zone d'interaction (comprise entre 0,05 et 0,1).

Caractérisation de la dimension moléculaire de l'IRES libre en solution par SAXS: la technique SAXS (Small Angle X-Rays Scattering), nous a permis de caractériser pour la première fois la dimension moléculaire de l'IRES libre en solution. Les données de SAXS seront utilisées afin de générer un modèle de l'IRES libre ou en complexe à une résolution quasi atomique. La résolution structurale de ce complexe RibonucleoProtéique impliqué dans la traduction du VHC aidera à mieux comprendre les phases initiales de la traduction initiale du VHC.



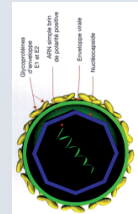
(A) : Etude structurale par SAXS du domaine II de l'IRES du VHC. Utilisation des programmes GASBOR et DAMMIN pour la génération de modèles expérimentaux qui sont superposés à la structure RMN du domaine II.
(B) : Etude structurale par SAXS du domaine II, III, IV de l'IRES du VHC. Utilisation du programme DAMMIN pour la génération de modèle expérimentaux. Données en cours de traitement.
(C) : Données structurales des modèles obtenus pour l'IRES dII, dIII, dIV par SAXS. Données en cours de traitement.

Conclusion & perspectives

L'étude structurale et fonctionnelle de ce complexe impliqué dans l'initiation de la traduction du VHC nous aidera à mieux comprendre les phases initiales du cycle viral. Les résultats que nous avons obtenu ici complètent les résultats publiés par d'autres équipes et serviront de base pour le « drug design » de molécules inhibant spécifiquement la traduction du VHC. C'est dans ce but que nous avons réussi à inhiber la formation de ce complexe à l'aide d'un anticorps dirigé contre le domaine eIF3b-MRR. Ce résultat est très encourageant dans le cadre du « drug discovery ». De plus les résultats obtenus en SAXS vont nous permettre de réaliser un modèle à haute résolution de l'IRES libre ou complexé avec différentes protéines (eIF3, eIF3b ou eIF3b-MRR) pour étudier finement les zones d'interactions mises en jeu lors de la formation du complexe eIF3/IRESHCV in vitro.

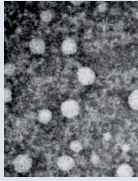
Human initiation factor eIF3 subunit b interacts with HCV IRES RNA through its N-terminal RNA recognition motif. FEBS Letters 583 (2009) 70-74.
Julien Pérard, Rodolfo Rasia, Jan Medenbach, Isabel Ayala, Jérôme Boisbouvier, Emmanuel Drouet, Florence Baudin.

Le virus de l'Hépatite C ?



Taille : 60nm de diamètre (1 million de fois plus petit qu'une balle de tennis).

Fonction : Parasite cellulaire qui détourne la machinerie cellulaire à son profit, c'est un **pirate !!**



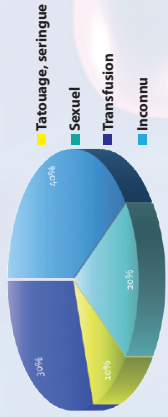
L'Hépatite C et les traitements

Type de maladie : infectieuse

Mode de transmission : sang, essentiellement lors de transfusion sanguine (année 80), tatouage...

Cible : foie

Evolution : Hépatite 80%, 20% cirrhose
Traitement : pas de vaccin, ralentissement des symptômes par l'utilisation d'anti-viraux peu efficaces.



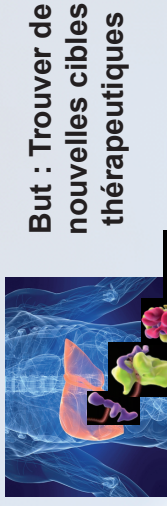
Le VHC : un problème de santé publique



Monde : 170 millions de personnes infectées soit 3% de la population.
France : 0,6 millions de personnes infectées
 60% sont des toxicoomanes
 25% des sidéens (1er cause de mortalité)
 15% suite à une transfusion

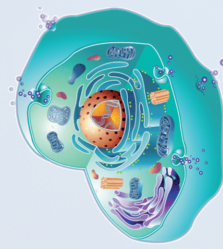
L'objectif de ma thèse

- Comment le virus fait-il pour pirater la cellule ?
- Comment stopper l'action du virus ?



But : Trouver de nouvelles cibles thérapeutiques

Détournement de la machinerie cellulaire par le virus



Le virus de l'Hépatite C, un pirate cellulaire

Julien Péraud, Allocataire de recherche 2006-2009, laboratoire UVHCI, directeur de thèse E. Douet et F. Baudin
Sujet de Thèse : Etude structurale et fonctionnelle de l'IRES de l'Hépatite C
Domaine de recherche : Virologie biologique moléculaire et biochimie des protéines

Les moyens mis en oeuvre

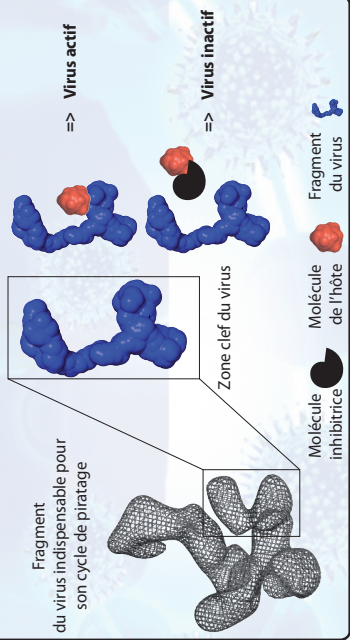
- Techniques :**
- Biologie moléculaire
 - Biochimie des protéines
 - Biologie cellulaire
 - Etudes structurales
 - Utilisation des grands instruments (ESRF)
 - Bio-informatique



- Collaboration :**
- IBS, UVHCI (Grenoble)
 - EMBL (Heidelberg)

Les résultats obtenus

- Identification d'un des mécanismes de piratage du virus
- Caractérisation de deux cibles pour une stratégie antivirale
- Identification d'une molécule qui perturbe le fonctionnement du virus
- Caractérisation de la forme globale d'une partie du génome viral



Une nouvelle piste pour une thérapie antivirale

Deux pistes à suivre :



- Cibler le virus
 Plus de piratages possible, le virus ne peut pas survivre

- Cibler l'hôte
 Protéger la cellule même si elle est infectée

Identification d'un mécanisme de piratage =

Cible thérapeutique

Résumé

Mots Clefs : IRES, VHC, MRR, eIF3, SAXS

L'IRES du VHC est organisé en une structure secondaire, complexe et hautement conservée, qui comprend quatre domaines distincts (I-IV). Cette organisation permet la liaison directe du facteur d'initiation eucaryote eIF3 et de la sous-unité 40S du ribosome. Le facteur d'initiation eucaryote eIF3 est un complexe de 13 protéines (~ 800 kDa). Certaines de ses sous-unités montrent une affinité pour l'ARN, et parmi elles deux possèdent des motifs de reconnaissance à l'ARN (MRR) : eIF3b (aa : 185-268) et eIF3g (aa : 239-317). Afin de déterminer l'implication de ces motifs dans l'interaction ARN-protéine, nous avons cloné les domaines contenant ces motifs MRR, ainsi que les protéines entières eIF3b et eIF3g. Nous avons ensuite étudié l'interaction de l'ARN et/ou domaines d'ARN avec le complexe eIF3 et/ou protéines ou domaines protéiques en utilisant la méthode de rétention sur filtre. Nous avons ainsi calculé les différentes constantes apparentes de dissociation entre : eIF3/IRES, eIF3b/IRES et MRR-eIF3b/IRES. Ces résultats montrent que la sous-unité eIF3b se lie directement au domaine III de l'IRES *via* son motif MRR situé dans le domaine N-terminal. Dans un deuxième temps, nous avons utilisé la spectroscopie RMN pour identifier précisément les acides aminés impliqués dans la reconnaissance de l'ARN viral. En utilisant différentes techniques telles que FBA, RMN et SAXS, il s'avère que le domaine IIIId (30 nucléotides) représente le motif minimum impliqué dans l'interaction MRR-eIF3b/IRES. Enfin, la technique de SAXS a été mise à profit pour étudier la structure tridimensionnelle de l'IRES libre en solution.

Abstract

Keywords : IRES, HCV, RRM, eIF3, SAXS

HCV IRES is organized in highly complex and conserved secondary structure, which includes four different domains (I-IV). This organization allows the direct binding of eucaryotik initiation factor 3 and of 40S ribosomal subunit. eIF3 is a complex of 13 proteins (~ 800 kDa). Some of subunies show an affinity for ARN, and among them two contain RNA Recognition Motif (RRM) : eIF3b (aa: 185-268) and eIF3g (aa: 239-317). To determine the implication of these motifs in the RNA/protein interaction, we cloned domains containing these RRM motifs, as well as whole proteins eIF3b and eIF3g. We then studied the interaction of RNA and/or RNA domains with the eIF3 complex and/or proteins or protein domains by using the method of Filter Binding assay. We calculated the various visible dissociation constants between: eIF3/IRES, eIF3b/IRES and MRR-eIF3b/RES. These results show that eIF3b sub-unit is directly bound to the domain III of the IRES *via* its motive RRM situated in the N-terminal domain. Secondly, we used NMR spectroscopy to identify exactly amino acids involved in the recognition of viral RNA. By using various techniques such as FBA, NMR and SAXS, it turns out between domain IIIId (30 nucléotides) represents the minimum motif involved in the interaction MRR-eIF3b/IRES. Finally, the technique of SAXS was used to study the three-dimensional structure of the free IRES in solution.