

THESE

Pour obtenir le titre de
DOCTEUR de l'ECOLE POLYTECHNIQUE

Discipline : Biologie

Présentée et soutenue publiquement par

Laurent Guillon

Le 7 juillet 2008

**Etude des facteurs du démarrage de la traduction
eIF3 et eIF5B**

Jury composé de :

Pr. Sylvain Blanquet	Président du Jury
Pr. Claudio Gualerzi	Rapporteur
Dr. Yves Mechulam	Directeur de thèse
Dr. Arnaud Poterszman	Examineur
Dr. Emmanuelle Schmitt	Codirectrice de thèse
Dr. Marat Yusupov	Rapporteur

RESUME

Le démarrage de la traduction est un processus central dans toute cellule. L'étude des protéines assistant le ribosome pour réaliser cette étape, les facteurs de démarrage (Initiation Factors Ifs), permet d'obtenir des informations sur les mécanismes moléculaires complexes assurant la fidélité et l'efficacité du démarrage.

La comparaison des jeux de facteurs protéiques dans les trois règnes du monde vivant a permis de mettre en évidence la présence de trois facteurs universellement conservés. Parmi ceux-ci, le facteur eucaryotique/archéen e/aIF5B, homologue au facteur bactérien IF2, stimule l'association des sous-unités ribosomales au même titre que chez les Bactéries. Néanmoins, l'universalité du facteur est limitée par l'absence d'interaction reportée entre le facteur e/aIF5B et l'ARNt initiateur alors que cette liaison est parfaitement caractérisée chez les Bactéries. Une partie de ce travail de thèse a permis d'étendre la similitude fonctionnelle entre les facteurs en mettant en évidence une liaison de l'ARNt initiateur méthionylé par le facteur e/aIF5B. Cette liaison présente des caractéristiques identiques à celle de l'ARNt initiateur méthionylé et formylé par le facteur bactérien IF2.

Une deuxième partie du travail de thèse a concerné le facteur eIF3, le plus complexe du système de démarrage chez les Eucaryotes. Ce complexe de 13 sous-unités chez l'humain et de 5 sous-unités chez la levure n'a pas d'équivalent dans les autres domaines du vivant bien qu'il joue un rôle central et essentiel chez les Eucaryotes. La compréhension de ses fonctions est néanmoins fortement limitée par le manque d'information à l'échelle moléculaire sur les interactions entre les sous-unités le composant et avec ses autres facteurs partenaires. De plus, le facteur s'avère être impliqué dans de nombreux cancers, ce qui étend l'intérêt de son étude. Mon travail a permis de développer une bibliothèque de vecteurs permettant de coexprimer les différentes sous-unités ou des formes stabilisées des sous-unités du facteur eIF3 de levure chez la Bactérie *Escherichia coli*. La purification des sous-unités isolées et de différents sous-complexes nous permet d'envisager la résolution de la structure du facteur et de son organisation par une approche alliant la cristallographie et la microscopie électronique.

ABSTRACT

Translation initiation plays a central role in every living organism. Studying the proteins helping the ribosome toward this task, Initiation Factors (Iifs), allows to get insights into the complex molecular mechanisms ensuring a high fidelity and efficiency of translation initiation.

When comparing the existing initiation factors in the three kingdoms of life, three universally conserved initiation factors were identified. Among these, the eukaryotic/archaeal factor e/aIF5B, the counterpart of the bacterial factor IF2, stimulates the ribosomal subunits joining step as in Bacteria. However, the universality of this factor was limited by the lack of reported interaction between the factor e/aIF5B and the initiator tRNA whereas such an interaction was characterized in Bacteria. A first part of this thesis work aimed at extending the functional similarity between the factors by bringing to light a binding of methionylated initiator tRNA by the factor e/aIF5B. The characteristics of this binding are very similar to the binding of formylated methionylated initiator tRNA by the bacterial factor IF2.

A second part of my thesis work dealt with the factor eIF3, the most complex factor of the eukaryotic translation initiation apparatus. This multimeric factor, composed of 13 subunits in human and 5 in budding yeast, has no counterpart in the other kingdoms of life despite a central and essential role in Eukaryotes. In addition, eIF3 is involved in numerous cancers. However, the understanding of its functions is severely limited by the lack of structural information on subunits interactions and on interactions with other partners of the translation apparatus. This work led to the development of a plasmid library allowing the coexpression of *Saccharomyces cerevisiae* eIF3's subunits or stabilized forms of them in the Bacteria *Escherichia coli*. The purifications of the individual subunits and different subcomplexes are leading us on the road of the factor structure determination by an approach combining crystallography and electron microscopy.

REMERCIEMENTS

En premier lieu, je remercie Sylvain Blanquet de m'avoir fait l'honneur de présider le jury de soutenance de cette thèse.

Mes remerciements s'adressent également à Claudio Gualerzi et Marat Yusupov qui ont eu le courage de rapporter mon travail, ainsi qu'à Arnaud Poterszman qui a bien voulu examiner cette thèse.

Je tiens à remercier Pierre Plateau pour m'avoir accueilli au sein du Laboratoire de Biochimie et permis de réaliser ce travail de thèse au sein d'une équipe dynamique et hautement compétente.

Ma plus profonde gratitude est destinée à mes directeurs de thèse, Yves Mechulam et Emmanuelle Schmitt, dont l'encadrement a largement contribué à la réussite de ce projet. Les échanges fructueux que nous avons nourris m'ont permis de bénéficier de leurs connaissances et de leurs compétences pointues en biologie. Enfin, l'investissement et la rigueur dans ce travail ne sont pas étrangers à l'atmosphère très conviviale de cet accompagnement.

Je souhaite adresser mes remerciements à l'équipe du Laboratoire de Chimie des Milieux Réactionnels de l'Ecole Polytechnique pour les nombreuses analyses réalisées en spectrométrie de masse, à l'équipe de Nicolas Boisset de l'Institut de Minéralogie et physique des Milieux Condensés de Paris (Campus Boucicaut) pour les enregistrements de clichés de microscopie électronique, ainsi que Gérard Keith de l'Institut de Génomique et de Biologie Structurale de Strasbourg pour son aimable contribution au projet.

J'adresse mes sincères remerciements à ma cothésarde Marie Naveau pour son écoute, sa gentillesse et les échanges profonds que nous avons eus tant au niveau professionnel que personnel.

Je remercie également les « jeun's », passés ou présents, du laboratoire pour les moments passés ensemble et nos « réunions informelles » : Caroline Aubard, Lionel Cladière, Laurent Giorgi, Guillaume Heugas, Guillaume Launay, Anne Lopes, Josselin Noirel, Laurent Mugherli, Marcel Schmidt Am Busch, Sandra Wydau et Laure Yatime sans oublier les bioinformaticiens Alexey Aleksandrov, Najette Amara, Damien Thompson et Pablo Tortosa.

Une pensée particulière est adressée aux femmes qui m'ont fourni une écoute et un réconfort auxquels je n'ai eu cesse d'avoir recours ; merci à Annick Carpentier, Laura Fioni, Catherine Morais et Catherine Teillet.

Je tiens à remercier mon « cofacteur nicotinique » Myriam Lazard pour nos échanges passionnants sur des sujets variés, son accompagnement, et l'amitié que nous partageons.

Je souhaite adresser plus généralement mes remerciements à tous les membres du laboratoire pour leur convivialité et leur dynamisme : mes voisins de paillasse, les « B2 » : Pierre-Damien Courreux, Christine Schurdevin et Michel Panvert ; mon voisin de bureau Marc Dauplais avec qui nous avons notamment partagé notre espace de travail (...) ; notre ACMO Michel Fromant pour les discussions très intéressantes et quelque fois animées (...) ; François Cluzeau, Pascal et Pascale Delplanche, Françoise Humbert, Alfonso Jaramillo, David Mignon, Mélanie Monguillon, Gérald Peyroche et Thomas Simonson.

Je remercie sincèrement Thierry Blancho pour sa patience et sa persévérance dans la lourde tâche de mise en forme de ce manuscrit.

Enfin, je tiens à remercier ma famille et mes proches pour leur patience et leur soutien durant ces années de thèse, et en particulier Sébastien.

SOMMAIRE

Table des figures, Table des tableaux, Liste des abreviations

Chapitre I Introduction	1
PARTIE A LE RIBOSOME, LE CŒUR DE LA MACHINERIE DU DEMARRAGE	5
I La biogénèse des ribosomes	5
II Le cœur universellement conservé du ribosome	7
III Différences entre ribosomes de diverses origines	17
PARTIE B LE DEMARRAGE DE LA TRADUCTION.....	22
I Le démarrage bactérien.....	23
II Le démarrage chez les Eucaryotes.....	26
III Le démarrage chez les Archées	29
PARTIE C RELATIONS STRUCTURE/FONCTION DES FACTEURS DU DEMARRAGE EUCARYOTIQUE.....	31
I Le complexe de prédémarrage 43S.....	31
II Recrutement de l'ARNm au complexe de prédémarrage 43S et balayage de l'ARNm...	39
III Sélection du codon de démarrage et réponse à l'appariement.....	45
IV Régulation & modes alternatifs du démarrage de la traduction	52
PARTIE D LE FACTEUR EIF3	58
I Composition d'eIF3 chez les Eucaryotes	58
II Organisation du facteur eIF3 de <i>S. cerevisiae</i>	60
III De l'organisation à la fonction.....	63
IV Le rôle d'eIF3 dans la régulation de la traduction et le cancer	67
CONCLUSION	70
Chapitre II Le facteur e/aIF5B	71
PARTIE A CLONAGE ET PURIFICATION DES FACTEURS DE DEMARRAGE E/AIF5B ET AIF1A.....	73
I Clonage	73
II Surproduction et purification	74
PARTIE B PROTECTION DU MET-ARNI INITIATEUR PAR E/AIF5B	75
I Le test de protection de la déacylation spontanée d'un ARNI aminoacylé.....	75
II Résultats.....	77
PARTIE C CRISTALLISATION DU FACTEUR	83
I Obtention des cristaux	83
II Diffraction.....	84
DISCUSSION.....	85

Chapitre III Le facteur eIF3.....	87
PARTIE A STRATEGIE DE PURIFICATION DU COMPLEXE EIF3 DE <i>SACCHAROMYCES CEREVISIAE</i>	88
I Clonage des gènes codant pour les sous-unités du facteur eIF3 de levure.....	89
II Essais de surproduction.....	90
III Mise en place d'une stratégie de coexpression	93
IV Clonages dans les vecteurs Duet et résultats préliminaires de la coexpression	96
V Implémentations de la stratégie de purification	101
VI Conclusion.....	103
PARTIE B PURIFICATIONS DE DIFFERENTES FORMES DU FACTEUR EIF3	104
I Stratégie d'expression	104
II Résultats des purifications	105
III Conclusion	110
PARTIE C LE COMPLEXE EIF3BC1-CC1-I-GC1, PREMIER PAS VERS LA RESOLUTION DE LA STRUCTURE DU FACTEUR EIF3.....	111
I Purification du complexe eIF3bC1-cC1-i-gC1.....	111
II Structure d'eIF3bC1-cC1-i-gC1 par microscopie électronique.....	113
III Vers une structure du facteur eIF3 par microscopie électronique.....	115
PARTIE D LES CRISTAUX D'EIF3I DE <i>SACCHAROMYCES CEREVISIAE</i> : UN PROBLEME CRISTALLOGRAPHIQUE	117
I Purification de la protéine	117
II Principe de la cristallisation.....	117
III Diffraction par les rayons X et acquisition des données.....	119
DISCUSSION.....	148
Chapitre IV Matériels et Méthodes	149
I Techniques générales de biologie moléculaire, réactifs et tampons de purification.....	150
II Etude du facteur e/aIF5B.....	158
III Le facteur eIF3.....	161
IV La sous-unité eIF3j, cristallisation.....	165
Annexes.....	170
Références Bibliographiques	188

SOMMAIRE DETAILLE

Table des figures, Table des tableaux, Liste des abreviations

Chapitre I Introduction	1
PARTIE A LE RIBOSOME, LE CŒUR DE LA MACHINERIE DU DEMARRAGE	5
I La biogénèse des ribosomes	5
II Le cœur universellement conservé du ribosome	7
1 La petite sous-unité	7
a Liaison de l'ARNt initiateur	8
b Liaison de l'ARNm et contrôle de l'appariement	10
2 La grande sous-unité	12
a L'ARNt au site P et le centre catalytique	12
b La protéine L1	13
c Le centre d'activation des GTPases	13
3 Liaison des sous-unités	14
a Liaison des sous-unités : la tête, un élément mobile	15
b Liaison des sous-unités : des réseaux d'interactions à la fonction	15
III Différences entre ribosomes de diverses origines	17
1 Les extensions de la petite sous-unité	17
a Les extensions de l'ARNr	17
b Les protéines ribosomales	18
2 Les extensions de la grande sous-unité	19
a Les extensions de l'ARNr	19
b Les protéines ribosomales	21
3 Les ponts entre les sous-unités	21
PARTIE B LE DEMARRAGE DE LA TRADUCTION	22
I Le démarrage bactérien	23
1 Liaison de l'ARNm	24
2 Liaison de l'ARNt initiateur méthionylé formylé	24
3 Le facteur IF2	24
4 Liaison d'IF2 et du fMet-ARNt _f ^{Met} au ribosome	25
5 Liaison d'IF1 au ribosome	25
6 L'ajustement du complexe formé	25
II Le démarrage chez les Eucaryotes	26
1 Liaison de l'ARNt initiateur : formation du complexe de prédémarrage 43S	26
2 Liaison de l'ARNm : le complexe eIF4F	28
3 La sélection du codon de démarrage : le balayage	28
4 L'association des sous-unités ribosomales : eIF5B	28
III Le démarrage chez les Archées	29
1 Liaison de l'ARNm : un démarrage bactérien	29
2 Formation d'un complexe 80S : un démarrage eucaryotique	29

PARTIE C RELATIONS STRUCTURE/FONCTION DES FACTEURS DU DEMARRAGE EUCARYOTIQUE.....	31
I Le complexe de prédémarrage 43S.....	31
1 Les composants du complexe 43S	31
a Le complexe ternaire eIF2:GTP:Met-ARNt _i ^{Met}	31
b Le facteur eIF5.....	34
c Le facteur eIF1A.....	35
d Le facteur eIF1.....	36
e Le facteur eIF3.....	36
2 Génération de sous-unités libres	37
3 Assemblage du complexe de prédémarrage 43S	38
II Recrutement de l'ARNm au complexe de prédémarrage 43S et balayage de l'ARNm...	39
1 Liaison de la coiffe en 5'	39
2 Liaison de la queue poly(A) en 3'	40
3 L'adaptateur eIF4G	41
4 L'hélicase eIF4A.....	42
5 Liaison du complexe eIF4F:ARNm au complexe 43S.....	43
6 Le balayage de l'ARNm	45
III Sélection du codon de démarrage et réponse à l'appariement.....	45
1 Sélection du codon de démarrage	45
2 Hydrolyse du GTP	46
3 Ejection du phosphate.....	47
4 Régénération du facteur eIF2 : eIF2B	48
5 L'association des sous-unités ribosomales : eIF5B	49
IV Régulation & modes alternatifs du démarrage de la traduction	52
1 Régulation de la liaison du Met-ARNt _i ^{Met} : phosphorylation d'eIF2 α	52
2 Régulation de la liaison de l'ARNm : phosphorylation d'eIF4E et 4E-BP	54
3 Régulation de la PABP	54
4 Démarrage par un Site Interne d'Entrée du Ribosome (IRES)	56
PARTIE D LE FACTEUR EIF3	58
I Composition d'eIF3 chez les Eucaryotes.....	58
II Organisation du facteur eIF3 de <i>S. cerevisiae</i>	60
III De l'organisation à la fonction	63
1 Le complexe Multi-Facteur (MFC) de <i>S. cerevisiae</i>	63
2 La formation du MFC stimule la formation du complexe 43S.....	65
IV Le rôle d'eIF3 dans la régulation de la traduction et le cancer	67
1 eIF3 et la régulation de la traduction	67
2 eIF3 et le cancer	68
CONCLUSION	70

Chapitre II Le facteur e/aIF5B	71
PARTIE A CLONAGE ET PURIFICATION DES FACTEURS DE DEMARRAGE E/AIF5B ET AIF1A.....	73
I Clonage	73
II Surproduction et purification	74
PARTIE B PROTECTION DU MET-ARNt INITIATEUR PAR E/AIF5B	75
I Le test de protection de la déacylation spontanée d'un ARNt aminoacylé.....	75
1 Principe du test	75
2 Mise en œuvre expérimentale.....	77
II Résultats.....	77
1 Protection de l'ARNt initiateur méthionylé par aIF5B	77
2 Compétition par l'ARNt _f ^{Met}	79
3 Spécificité de la liaison	79
4 Cartographie de l'interaction	81
5 Influence de la liaison à aIF1A	82
6 Protection de l'ARNt initiateur méthionylé par eIF5B	82
PARTIE C CRISTALLISATION DU FACTEUR	83
I Obtention des cristaux	83
II Diffraction.....	84
DISCUSSION.....	85
 Chapitre III Le facteur eIF3.....	 87
PARTIE A STRATEGIE DE PURIFICATION DU COMPLEXE EIF3 DE SACCHAROMYCES CEREVISIAE	88
I Clonage des gènes codant pour les sous-unités du facteur eIF3 de levure.....	89
II Essais de surproduction.....	90
III Mise en place d'une stratégie de coexpression	93
IV Clonages dans les vecteurs Duet et résultats préliminaires de la coexpression	96
1 Clonages	96
2 Purification du complexe eIF3 dans le système de coexpression.....	97
3 Caractérisation d'eIF3g	99
V Implémentations de la stratégie de purification	101
1 Clonages	101
2 Caractérisation d'eIF3bC1	101
3 Stœchiométrie des complexes.....	102
VI Conclusion.....	103

PARTIE B PURIFICATIONS DE DIFFERENTES FORMES DU FACTEUR EIF3	104
I Stratégie d'expression	104
1 Expression des protéines	104
2 Purification	104
II Résultats des purifications	105
1 Les sous-unités isolées	105
2 Les complexes à deux partenaires	105
3 Les complexes trimériques et tétramériques.....	105
III Conclusion	110
PARTIE C LE COMPLEXE EIF3BC1-CC1-I-GC1, PREMIER PAS VERS LA RESOLUTION DE LA STRUCTURE DU FACTEUR EIF3.....	111
I Purification du complexe eIF3bC1-cC1-i-gC1	111
II Structure d'eIF3bC1-cC1-i-gC1 par microscopie électronique.....	113
III Vers une structure du facteur eIF3 par microscopie électronique.....	115
PARTIE D LES CRISTAUX D'EIF3I DE <i>SACCHAROMYCES CERVISIAE</i> : UN PROBLEME CRISTALLOGRAPHIQUE	117
I Purification de la protéine	117
II Principe de la cristallisation.....	117
1 Principe	117
2 Cristallisation d'eIF3i	118
III Diffraction par les rayons X et acquisition des données.....	119
1 La diffraction des rayons X par les cristaux et conditions de diffraction.....	119
2 Source du rayonnement X et cryoprotection.....	122
3 Collecte et traitement des données.....	123
4 Analyses cristallographiques préliminaires	126
a Contenu de la maille	126
b Symétries non cristallographiques, la fonction de Patterson.....	127
c Détermination du groupe d'espace correct et contrôle des données	128
d Emulation de la maille I222 par la maille C2	128
e Jumelage : origine et conséquence.....	130
5 Phasage des données	134
a Phasage par remplacement moléculaire	134
b Phasage par MIR.....	139
c Phasage par MAD.....	143
DISCUSSION.....	148

Chapitre IV Matériels et Méthodes 149

I Techniques générales de biologie moléculaire, réactifs et tampons de purification.....	150
1 Souches de bactéries.....	150
2 Plasmides, préparations plasmidiques et caractérisation	151
3 Clonage.....	151
4 Mutagenèse dirigée.....	153
5 Séquençage de l'ADN.....	154
6 Caractérisation des protéines par SDS-PAGE.....	154
7 Caractérisation des protéines par électrophorèse bidimensionnelle.....	154
8 Caractérisation des protéines par électrophorèse en condition native.....	154
9 Cocktail antiprotéases.....	155
10 Purification à grande échelle des protéines	156
11 Essais de surproduction des protéines comportant une étiquette His ₆	157
12 Digestion de bandes excisées de gels pour analyse par SM.....	157
13 Autres produits et réactifs	157
II Etude du facteur e/αIF5B.....	158
1 Clonage et Purification de ss-αIF5B.....	158
2 Clonage et Purification des domaines de ss-αIF5B	158
3 Purification de ss-αIF1A.....	159
4 Clonage et purification du facteur pa-αIF5B.....	159
5 Clonage et Purification de γ-ΔeIF5B	159
6 Purification des ARNt.....	160
7 Tests de protection.....	160
8 Cristallisation et cryoprotection	161
III Le facteur eIF3.....	161
1 Clonages pour obtenir la bibliothèque de vecteurs	161
2 Expression et purification	161
3 Les sous-unités isolées.....	162
a eIF3aN1, eIF3aN2 (pET15bXSB-eIF3aN1, pET15bXSB-eIF3aN2).....	162
b eIF3bC1 (pET32b-eIF3bC1).....	162
c eIF3gC1 (pET15bIpa-eIF3gC1)	163
d eIF3i (pET15bIpa-eIF3i).....	163
4 Purifications de complexes à deux partenaires.....	163
a eIF3bC2-3cC1 (pET32b-eIF3bC2 + pCDF-3cC1).....	163
b eIF3i-3gC1 (pCOLA-3i-3gC1).....	163
c eIF3a-3cC1 (pCDF-3a-3cC1).....	163
d eIF3bC1-3i (pET32b-3bC1 + pCOLA-3i)	164
e eIF3bC1-3a (pET32b-3bC1 + pCDF-3a)	164
f eIF3cC1-3i (pET15bIpa-3i + pCDF-3cC1)	164

5 Purifications de complexes trimériques.....	164
a eIF3bC1-3i-3gC1 (pET32b-3bC1 + pCOLA-3i-3gC1)	164
b eIF3a/3aN1/3aN2-3bC1-3cC1 (pET32b-3bC1 + pCDF-3a/3aN1/3aN2-3cC1)	164
c eIF3cC1-3i-3gC1 (pET15bIpa-eIF3cC1/pCDF-3cC1 + pCOLA-3i-3gC1).....	164
6 Purifications de complexes tétramériques.....	165
a eIF3a-3cC1-3i-3gC1 (pCDF-3a-3cC1 + pCOLA-3i-3gC1).....	165
b eIF3bC1-cC1-i-gC1.....	165
7 Purifications du complexe pentamérique	165
IV La sous-unité eIF3i, cristallisation.....	165
1 Cristallisation	165
2 Expression et purification des protéines sélénométhionylées.....	166
3 Production d'eIF3i sous forme native	166
4 Production d'eIF3i fusionnée à une étiquette C-terminale	166
5 Production des mutants de délétion des résidus C-terminaux.....	167
6 Obtention des mutants cystéines multiples.....	167
Annexes.....	170
Références Bibliographiques.....	188

TABLE DES FIGURES

Figure 1.1 :	Assemblage, maturation et export des sous-unités ribosomales 40S et 60S eucaryotiques.....	6
Figure 1.2 :	Vue de la sous-unité 30S du ribosome bactérien depuis son interface avec la grande sous-unité .	7
Figure 1.3 :	Représentation en « feuille de trèfle » et structure tridimensionnelle de l'ARNt	8
Figure 1.4 :	Liaison du bras anticodon de l'ARNt au site P.....	9
Figure 1.5 :	Contrôle de l'appariement codon-anticodon aux sites A et P	10
Figure 1.6 :	Vue de la sous-unité 50S du ribosome bactérien depuis son interface avec la petite sous-unité .	12
Figure 1.7 :	Vue du centre peptidyl-transférase de la sous-unité 50S.....	13
Figure 1.8 :	Vue des ponts entre les sous-unités 30S et 50S.....	15
Figure 1.9 :	Partition ARN/Protéines ribosomales de la sous-unité 40S de <i>S. cerevisiae</i>	18
Figure 1.10 :	Vues des grandes sous-unités ribosomales de différents organismes	20
Figure 1.11 :	Ponts entre les sous-unités 30S (jaune) et 50S (bleu) vues depuis leurs domaines de liaison.....	21
Figure 1.12 :	Le démarrage de la traduction chez les Bactéries	23
Figure 1.13 :	Le démarrage de la traduction eucaryotique	27
Figure 1.14 :	Domaines importants des facteurs eIF2 β (numérotation de <i>S. cerevisiae</i>) et aIF2 β (numérotation <i>Pyrococcus abyssi</i>)	31
Figure 1.15 :	Structure tridimensionnelle d'aIF2 γ de <i>P. abyssi</i>	32
Figure 1.16 :	Structure tridimensionnelle d'eIF2 α de <i>P. abyssi</i>	33
Figure 1.17 :	Structure de l'hétérotrimère eIF2 de <i>Sulfolobus solfataricus</i> et comparaison au dimère $\beta\gamma$ de <i>P. abyssi</i>	33
Figure 1.18 :	Domaines importants du facteur eIF5.....	34
Figure 1.19 :	Structure tridimensionnelle du domaine C-terminal d'eIF5.....	34
Figure 1.20 :	Conservation du facteur IF1/e/aIF1A et structure du facteur humain.....	35
Figure 1.21 :	Homologie structurale entre les facteurs eIF1 humain et IF3 bactérien	36
Figure 1.22 :	Structure de la sous-unité 40S liée à eIF3 et eIF4G	37
Figure 1.23 :	Structure de la sous-unité 30S liée à IF1 et IF3	38
Figure 1.24 :	Liaison de la coiffe par le facteur eIF4.....	39
Figure 1.25 :	Liaison d'un polymère d'adénine par les RRM1 et 2 de la PAB	40
Figure 1.26 :	Structure tridimensionnelle du complexe eIF4E:m ⁷ GTP:eIF4GI	41
Figure 1.27 :	Visualisation de complexe ARN/protéines circulaires par microscopie à force atomique	42
Figure 1.28 :	Modèle du mécanisme d'action de la portion centrale d'eIF4G sur la configuration des domaines d'eIF4A.....	43
Figure 1.29 :	Schéma des changements de conformations de la sous-unité 40S par la liaison des facteurs de démarrage.....	44
Figure 1.30 :	Modèle de reconnaissance de l'appariement codon-anticodon	47
Figure 1.31 :	Domaines importants du facteur eIF2B ϵ	48
Figure 1.32 :	Structure du facteur aIF5B de <i>Methanobacterium thermoautotrophicum</i>	49
Figure 1.33 :	Visualisation du mouvement des domaines d'aIF5B induit par la liaison du GTP-Mg ²⁺	50
Figure 1.34 :	Comparaison des états liés au GTP et au GDP d'IF2	51
Figure 1.35 :	Régulation de la traduction de <i>GCN4</i> et <i>ATF4</i>	53
Figure 1.36 :	Rôle de la PABP et la queue poly(A) dans la traduction	55
Figure 1.37 :	Mode d'action des IRES du HCV et du CrPV	57
Figure 1.38 :	Principe du double hybride avec le système Gal4 UAS-LacZ.....	60
Figure 1.39 :	Représentation schématique des résultats obtenus.....	61

Figure 1.40 :	Représentation schématique des résultats obtenus.....	62
Figure 1.41 :	Représentation schématique de l'ensemble des résultats obtenus sur les interactions entre les sous-unités du facteur eIF3	63
Figure 1.42 :	Organisation du MFC par le facteur de démarrage eIF3.....	64
Figure 1.43 :	Modèle proposé par Alan Hinnebusch et al.....	66
Figure 2.1 :	Analyse SDS-PAGE (12%) des protéines ss-eIF5B (A-piste 2) ou γ - Δ eIF5B (B-piste 2)	74
Figure 2.2 :	Détermination de la constante de dissociation (Kd) du complexe entre ss-eIF5B et le Met-ARNt _i ^{Met} de <i>S. solfataricus</i>	78
Figure 2.3 :	Spécificité de la liaison de l'ARNt aminoacylé par ss-eIF5B	80
Figure 2.4 :	Spécificité de la liaison de l'ARNt aminoacylé par le domaine IV de ss-eIF5B.....	81
Figure 2.5 :	Structure du facteur eIF5B de <i>Methanobacterium thermoautotrophicum</i>	85
Figure 3.1 :	Schéma de la démarche suivie pour la résolution de la structure du facteur de démarrage eIF3 ..	88
Figure 3.2 :	Le plasmide pET15blpa	89
Figure 3.3 :	Profils d'élution de la résine TALON des sous-unités isolées	91
Figure 3.4 :	Profil de la purification réalisée par co-cassage des 5 sous-unités.....	92
Figure 3.5a :	Organisation du MFC par le facteur de démarrage eIF3.....	94
Figure 3.5b :	Représentation schématique de la sous-unité eIF3b et des différentes troncutures envisagées ..	95
Figure 3.6 :	Représentation schématique de la sous-unité eIF3c et des différentes troncutures envisagées ..	96
Figure 3.7 :	Représentation schématique des clonages réalisés dans le système Duet	97
Figure 3.8 :	Profil de la purification par coexpression des 5 sous-unités.....	98
Figure 3.9 :	Effet de l'ajout d'EDTA sur la migration d'eIF3g sur gel SDS-PAGE	99
Figure 3.10 :	Représentation schématique des sous-unités eIF3a et eIF3g et des différentes troncutures envisagées.....	101
Figure 3.11 :	Représentation des deux systèmes de coexpression (A et B) envisagés.....	103
Figure 3.12 :	Analyse par électrophorèse (SDS-PAGE 12%) du complexe eIF3bC1-cC1-i-gC1 purifié.....	111
Figure 3.13 :	Analyse des deux pics de l'hétérotétramère eIF3bC1-cC1-i-gC1	113
Figure 3.14 :	Images de microscopie électronique de l'hétérotétramère eIF3bC1-cC1-i-gC1	114
Figure 3.15 :	Cristaux d'eIF3i obtenus dans la condition 1,8M Na/KPO4 pH8,2	119
Figure 3.16 :	Cliché de diffraction d'un cristal d'eIF3i natif	124
Figure 3.17 :	Représentation des mailles cristallographiques I222 et C2.....	129
Figure 3.18 :	Représentation de la maille cristallographique C2 et de sa maille jumelle	133
Figure 3.19 :	Structure tridimensionnelle de WDR5.....	136
Figure 3.20 :	Alignement de séquences de β -propeller et d'eIF3i réalisé avec Clustal W	137
Figure 3.21 :	Visualisation de la détermination des phases par MIR.....	139
Figure 3.22a :	Alignement de séquences d'eIF3i réalisé avec PipeAlign.....	146
Figure 3.22b :	Alignement de séquences d'eIF3i réalisé avec PipeAlign.....	147

TABLE DES TABLEAUX

Tableau 1.1 : Facteurs du démarrage de la traduction dans les trois domaines du vivant	22
Tableau 1.2 : Nomenclature unifiée des sous-unités du facteur eIF3	59
Tableau 2.1 : Influence de la nature de l'ARNt initiateur sur sa liaison par les facteurs ss-eIF5B et pa-eIF5B ...	78
Tableau 2.2 : Influence du nucléotide lié à ss-eIF5B sur sa liaison à l'ARNt initiateur d' <i>E. coli</i>	79
Tableau 2.3 : Liaison du Met-ARNt _f ^{Met} aux domaines de ss-eIF5B.....	81
Tableau 3.1 : Analyse en spectrométrie de flamme du contenu en zinc de la protéine eIF3gC1	100
Tableau 3.2 : Bilan des purifications des sous-unités isolées	107
Tableau 3.3 : Bilan des purifications des dimères	108
Tableau 3.4 : Bilan des purifications des trimères et tétramères.....	109
Tableau 3.5 : Atomes lourds testés pour le phasage des cristaux d'eIF3i par la méthode MIR	142
Tableau 4.1 : Génotype des souches utilisées dans cette étude	150
Tableau 4.2 : Concentrations d'antibiotique utilisées pour les sélections	151
Tableau 4.3 : Composition des tampons utilisés pour l'électrophorèse	155
Tableau 4.4 : Composition du cocktail antiprotéase CAPx50.....	155
Tableau 4.7 : Oligonucléotides pour produire les différents plasmides permettant de produire eIF3i	167
Tableau 4.5 : Bibliothèque de vecteurs.....	168
Tableau 4.6: Oligonucléotides utilisés pour obtenir la bibliothèque de vecteurs	169

LISTE DES ABREVIATIONS

ADN	Acide DésoxyriboNucléique
aIF	Facteur archéen du démarrage de la traduction
Amp	Ampicilline
ARN	Acide RiboNucléique
ARNm	ARN messenger
ARNr	ARN ribosomal
ARNt	ARN de transfert
ATP	Adénosine TriPhosphate
Cm	Chloramphénicol
DO	Densité Optique
DTT	DiThioThrétol
EDTA	EthylèneDiamine-tétra-Acétate
EF1A	Facteur d'Elongation bactérien 1A (ou Ef-Tu)
eIF	Facteur eucaryotique de démarrage de la traduction
ESRF	European Synchrotron Radiation Facilities
fMet-ARN _t ^{fMet}	ARNt initiateur bactérien méthionylé et formylé
GAP	GTPase Activating Protein
GDPNP	5'-Guanylylimido-DiPhosphate
GEF	Guanine Exchange Factor
GTP/GDP	Guanosine Tri/DiPhosphate
HEPES	Acide 4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonique
IF1, IF2, IF3	Facteurs bactériens du démarrage de la traduction 1, 2 et 3
IPTG	IsoPropyl-β-D-ThioGalactoside
IRES	Internal Ribosome Entry Site
Kan	Kanamycine
Met-ARN _t ^{Met}	ARNt initiateur eucaryotique ou archéen méthionylé
MFC	Multi-Factor Complex
MOPS	Acide 3-[N-Morpholino]PropaneSulfonique
OB fold	Oligonucleotide/Oligosaccharide Binding fold
PABP	Poly(A) Binding Protein
PCR	Polymerase Chain Reaction
PEG	PolyEthylène Glycol
Pi	Phosphate inorganique
PMSF	PhenylMethylSulfonyl Fluoride
PTC	Peptidyl-Transferase Centre
RMN	Résonance Magnétique Nucléaire
SD	Séquence de Shine-Dalgarno
SDS	DodécylSulfate de Sodium
SDS-PAGE	Electrophorèse sur Gel de PolyAcrylamide en présence de SDS
Spec	Spectinomycine
Stre	Streptomycine
Tris	Tris(hydroxyméthyl)aminométhane

Chapitre I Introduction

Tous les organismes du monde vivant conservent leurs patrimoines génétiques sous une forme codée et stable : l'acide désoxyribonucléique (ADN). Les processus aboutissant au décodage de cette information doivent donc être hautement contrôlés afin d'assurer l'intégrité de l'organisme. Une analyse des gènes impliqués dans le transfert de l'expression génétique révèle une très grande conservation dans tous les domaines de la vie. En fait, ceux-ci feraient partie de «gènes de cœur» hérités d'un ancêtre universel commun ([Harris et al., 2003](#)). Les étapes qui assurent l'expression génétique s'intègrent dans deux grands événements du cycle cellulaire, universellement conservés, que sont la transcription et la traduction.

La transcription vise à extraire de l'ADN des portions nécessaires à un moment donné au fonctionnement de la cellule. Une enzyme, l'ARN polymérase assure la biosynthèse d'un polymère d'acide ribonucléique (ARN) complémentaire du fragment d'ADN d'intérêt. Cette complémentarité est identique à celle observée au sein de « l'alphabet à 4 lettres » qu'utilise l'ADN et qui lui confère sa structure bicaténaire (appariements Watson-Crick : A:T / C:G). Des différences notables existent entre ces deux biopolymères et l'ARN apparaît comme un support moins stable pour conserver l'information génétique. En premier lieu, les bases de l'ARN sont attachées sur des riboses (plus sensibles à l'hydrolyse que les désoxyriboses de l'ADN). La base Thymine (T) de l'ADN est substituée par une base Uracile (U) toujours complémentaire de la base Adénine (A), moins coûteuse à produire pour la cellule mais lentement convertie en cytosine (C). Enfin, les molécules d'ARN sont plus courtes que l'ADN et le plus souvent simple brin.

Les ARN synthétisés peuvent soit directement participer à l'activité cellulaire -tels l'ARN ribosomal (ARNr), les ARN de transfert (ARNt), les petits ARN nucléaires...-, soit être utilisés comme matrices lors de la deuxième phase de l'expression génétique. Cette dernière catégorie comprend les ARN messenger (ARNm) contenant les gènes qui seront ensuite traduits en protéines. Alors que les ARNm procaryotiques peuvent contenir plusieurs gènes -ils sont dits polycistroniques-, les ARNm eucaryotiques sont le plus souvent monocistroniques. De plus, les ARNm des Eucaryotes résultent de la modification d'un ARN dit pré-messenger. Au cours de la transcription, des enzymes ajoutent à l'extrémité 5' de l'ARN pré-messenger une structure particulière appelée coiffe. Cette dernière stimule toute une chaîne de réactions afin de produire un ARNm mature, à savoir épissé et doté d'une extrémité 3' polyadénylée (polyA). L'épissage permet l'excision de régions non codantes (ou *introns*) afin de rassembler les régions codantes (ou *exons*). En outre, la coiffe en 5' et la queue polyA en 3' stimulent l'export des ARNm matures du noyau vers le cytoplasme, lieu de la deuxième phase de l'expression génétique ([Lewis and Izaurralde, 1997](#)). En revanche, les deux phases de l'expression génétique procaryotique sont réalisées simultanément.

La deuxième étape, la traduction, assure la biosynthèse des protéines à partir des ARNm. Il s'agit de faire correspondre à chaque triplet de base de l'ARNm (ou codon) un acide aminé. De même que les acides nucléiques, les protéines sont dotées d'un « alphabet », mais à 20 lettres ; les acides aminés. La correspondance entre codons et acides aminés constitue le code génétique. Cette correspondance est assurée par un ensemble de petites molécules d'ARN fonctionnels, les ARNt. Ceux-ci présentent à une extrémité un triplet de base complémentaire d'un codon (ou anticodon) et portent à l'autre extrémité l'acide aminé correspondant au codon. L'accrochage de l'acide aminé sur l'ARNt est réalisé par une famille d'enzymes hautement spécifiques de leurs substrats, les aminoacyl-ARNt synthétases. Il existe cependant plus de codons (64) que d'acides aminés (20), plusieurs codons vont alors être associés à un même acide aminé. Ceci induit une redondance du code génétique mais il n'existe pas d'ambiguïtés : un codon ne code que pour un seul acide aminé. La synthèse des protéines se déroule au niveau d'une molécule ribonucléoprotéique, le ribosome, obtenu par le repliement des ARNr sur eux-mêmes et l'accrochage conjoint de nombreuses protéines dites ribosomales. Ces macromolécules sont qualifiées de ribozymes, elles possèdent une activité enzymatique propre grâce à leur partie ARN.

Schématiquement, au niveau d'un codon de l'ARNm, un ARNt portant l'anticodon correspondant se lie au ribosome. L'acide aminé porté par l'ARNt apparié sur le codon précédent se trouve alors à proximité de celui porté par l'ARNt nouvellement arrivé. La réaction de polymérisation entre les acides aminés s'effectue. De proche en proche, la chaîne polypeptidique est allongée au fur et à mesure que l'ARNm est traduit. A l'instar de toutes les réactions de polymérisation, la traduction se décompose en trois étapes : le démarrage, l'allongement et la terminaison. Cette succession d'étapes et le caractère impératif de haute fidélité témoignent de la complexité inhérente au processus de traduction ; ainsi, le ribosome requiert l'assistance d'un grand nombre de protéines afin d'optimiser son fonctionnement.

Le démarrage de la traduction, qui nous intéressera plus particulièrement, constitue l'étape critique et limitante de la biosynthèse des protéines. En effet, au cours de cette étape, le codon à partir duquel la production du polypeptide s'amorce est sélectionné. Or, le décalage d'une base ou le choix d'un mauvais codon de démarrage auraient des conséquences dramatiques sur la synthèse de la protéine. Ainsi, le choix de la phase pour la « lecture » -ou cadre de lecture- de l'ARNm (ORF : Open Reading Frame), c'est-à-dire la sélection du codon de démarrage et donc des codons qui lui font suite, représente la difficulté majeure à laquelle la machinerie de la traduction doit faire face. Ce choix est facilité par l'existence d'un codon universel de démarrage de la traduction (AUG) qui conditionne l'incorporation d'une méthionine (ou d'une méthionine formylée dans le cas bactérien) en premier acide aminé de toutes les protéines synthétisées. De plus, le codon de démarrage est décodé par un ARNt méthionine spécifique dit initiateur qui ne sera pas utilisé dans la phase d'allongement.

Bien qu'il puisse être décrit un schéma général pour décrire l'étape de la traduction pour tous les organismes, quelques différences nous sont déjà apparues entre les Bactéries, les Archées et les Eucaryotes. Celles-ci se trouvent renforcées si l'on s'intéresse un peu plus au jeu de facteurs protéiques assistant le ribosome (ou facteurs de démarrage) et aux différentes étapes élémentaires du démarrage de la traduction. Ainsi, dans tous les domaines du vivant, un facteur de démarrage caractéristique assure le recrutement de l'ARNt initiateur au ribosome. En effet, le facteur hétérotrimérique eucaryotique/archéen IF2 (e/aIF2), qui amène l'ARNt initiateur au ribosome, ne possède pas d'homologie autre que fonctionnelle avec le facteur bactérien IF2 chargé de transporter l'ARNt initiateur portant la méthionine formylée. Une analyse des génomes a cependant révélé l'existence tout à fait inattendue d'un facteur de démarrage eucaryotique et archéen fortement homologue au facteur IF2 bactérien : e/aIF5B. e/aIF5B est en fait impliqué dans une phase ultérieure du démarrage de la traduction. La présence des facteurs e/aIF2 et e/aIF5B pour se substituer à la fonction du facteur IF2 conduit à s'interroger sur les raisons de la coexistence de ces deux facteurs. Finalement, le jeu de facteurs de démarrage de la traduction archéenne ressemble plus à celui du système eucaryotique qu'à celui bactérien. Or, certaines protéines sont exclusivement présentes au sein de la machinerie eucaryotique. L'un des exemples les plus frappants concerne le facteur multimérique eIF3, le plus gros facteur intervenant lors du démarrage. Les études font de plus en plus état d'un rôle central d'eIF3 au niveau de la traduction et on lui associe notamment un rôle clé dans la régulation de la biosynthèse des protéines. A nouveau, nous sommes amenés à nous demander pourquoi une protéine si importante pour la traduction chez les Eucaryotes existe uniquement chez ces organismes. Conceptuellement, l'étude de ces protéines permet de mieux caractériser les étapes élémentaires du démarrage de l'organisme étudié et ainsi proposer un schéma d'évolution pour chaque domaine du vivant. De plus, les dérèglements de la machinerie traductionnelle sont souvent liés à des pathologies. L'implication récente du facteur eIF3 dans de nombreux types de cancer le désigne comme une cible thérapeutique potentielle. Ainsi, la caractérisation de ce facteur revêt une grande importance à de multiples niveaux.

Cette introduction traitera principalement du démarrage de la traduction chez les Eucaryotes. Une description de l'organe de cœur de la traduction, le ribosome, sera tout d'abord proposée. Ensuite, un aperçu des systèmes de démarrage bactérien et archéen permettra de souligner les divergences et les ressemblances existantes. Il s'en suivra une description plus détaillée des facteurs du démarrage chez les Eucaryotes et notamment leurs modes d'action dans ce processus. Enfin, nous nous intéresserons plus particulièrement au facteur de démarrage eIF3, qui constitue une partie de ce travail de thèse.

PARTIE A LE RIBOSOME, LE CŒUR DE LA MACHINERIE DU DEMARRAGE

Un nombre sans cesse croissant de données structurales obtenues par cristallographie et diffraction aux rayons X combinées aux résultats fournis par la cryomicroscopie électronique nous permettent de visualiser des éléments isolés et des complexes intervenant au cours du démarrage de la traduction dans les trois domaines du vivant. Notre connaissance à l'échelle moléculaire des mécanismes élémentaires s'en trouve très largement précisée. Néanmoins, progresser vers un aperçu dynamique du processus de démarrage nécessite que les efforts se concentrent sur le piégeage d'intermédiaires réactionnels afin de compléter les études biochimiques existantes. Dans cette introduction, nous débuterons par une description de l'organe au cœur de la machinerie du démarrage, à savoir le ribosome.

I La biogénèse des ribosomes

La biogénèse des particules ribonucléoprotéiques que sont les ribosomes est dans tous les organismes un processus coûteux tant au niveau de l'énergie que des métabolites requis. Chez la levure *Saccharomyces cerevisiae*, la transcription de l'ARNr représente environ 60% de la transcription totale et celle des ARNm des protéines ribosomales requiert environ 50% de l'activité totale de l'ARN polymérase spécifique de la traduction des ARNm (Pol II) ([Warner, 1999](#)).

Les ribosomes se composent de deux sous-unités (l'une dite petite et l'autre grande sous-unité) nommées en fonction de leurs coefficients de sédimentation (exprimés en Svedberg S). Le repliement de l'ARNr crée une structure secondaire comportant comme éléments architecturaux de nombreuses structures de type tige-boucle. Le maintien de la structure tertiaire fait intervenir de nombreuses interactions ARN-ARN, des interactions ARN-protéines et des interactions à plus longue distance pontées par des ions magnésiums stabilisant les ARN ([Ban et al., 2000](#)).

Les ribosomes des Bactéries et des Archées sédimentent à 70S pour une masse d'environ 2,5MDa. Ils comprennent une petite sous-unité 30S et une grande sous-unité 50S. Grossièrement, les deux tiers du ribosome sont composés d'ARN et les protéines ribosomales représentent le dernier tiers ([Schuwirth et al., 2005](#)). Chez *Escherichia coli*, la petite sous-unité comprend l'ARNr 16S (1542 nucléotides) et 21 protéines ribosomales alors que la grande sous-unité est composée des ARNr 23S (2904 nt) et 5S (120 nt) et de 33 protéines. La biogénèse débute par la synthèse d'un transcrite primaire comprenant les ARNr 16S, 23S et 5S. Au cours de la transcription, l'ARNr est mûré par diverses modifications. Simultanément, de nombreuses RNases génèrent les ARNr isolés qui se replient alors sur eux-mêmes. Dès que des sites de liaison de protéines ribosomales apparaissent, ces dernières viennent s'y ancrer. La formation de la sous-unité 50S par association des ARNr 23S et 5S liés aux protéines ribosomales demeure un processus complexe. Les repliements des sous-unités peuvent néanmoins être reproduits *in vitro* par mélange des ARNr et des protéines ribosomales, suivant des protocoles complexes ([Kaczanowska and Ryden-Aulin, 2007](#)).

Chez les Eucaryotes, le repliement des sous-unités ribosomales requiert l'intervention de nombreux facteurs protéiques et ARN non ribosomaux. Les cellules eucaryotiques doivent en effet assembler plus de 70 protéines ribosomales avec 4 ARNr différents. Les ribosomes eucaryotiques 80S, d'une masse avoisinant les 4MDa, se composent d'une petite sous-unité 40S et d'une grande sous-unité 60S. La sous-unité 40S de *Saccharomyces cerevisiae* compte l'ARNr 18S (1798 nt) et 32 protéines ribosomales. La sous-unité 60S est quant à elle composée des ARNr 25S (3392 nt), 5,8S (158 nt), 5S (121 nt) et de 45 protéines. La synthèse des ARNr et des ARNm est soumise à l'action concertée des trois machineries de la transcription par un mécanisme en plusieurs étapes illustré dans la **Figure 1.1**. Leur parfaite coordination conditionne dès lors l'efficacité et la fiabilité de la production de ribosome. L'ARN polymérase I (Pol I) permet la synthèse d'un transcrite primaire 35S. Pol III assure la transcription de l'ARNr 5S alors que Pol II produit les ARNm codant

les protéines ribosomales. Les pré-ARNr 35S et 5S ainsi que les protéines ribosomales, des protéines non ribosomales et des petits ARN nucléaires s'associent pour former une particule dite pré-ribosomale pré-90S. Une coupure au niveau du pré-ARNr 35S dissocie ensuite la particule pré-90S en précurseur de chacune des sous-unités : pré-60S (comprenant l'ARNr 5S et le pré-ARNr 27S) et pré-40S (pré-ARNr 20S). Ces pré-ribosomes sont ensuite exportés dans le cytoplasme, lieu de la traduction, via des pores nucléaires. De façon concomitante, les protéines non-ribosomales se dissocient et la maturation du pré-ARNr 27S en ARNr 25S et 5,8S crée la grande sous-unité 60S et celle du pré-ARNr 20S en ARNr 18S forme la petite sous-unité 40S (Tschochner and Hurt, 2003).

La description des sous-unités chez *E. coli* et *S. cerevisiae* fait apparaître des différences notables : la petite sous-unité ribosomale eucaryotique comprend 256 nucléotides et 11 protéines additionnels alors que sa grande sous-unité présente 646 nucléotides et 12 protéines supplémentaires. Les structures des ribosomes eucaryotiques obtenues par cryomicroscopie électronique, combinées aux structures déterminées par cristallographie et diffraction aux rayons X des ribosomes bactériens et archéens, ont permis de préciser ces différences. Intéressons nous tout d'abord à l'unité existant entre les ribosomes des différents organismes en nous focalisant davantage sur les éléments déterminants du démarrage de la traduction.

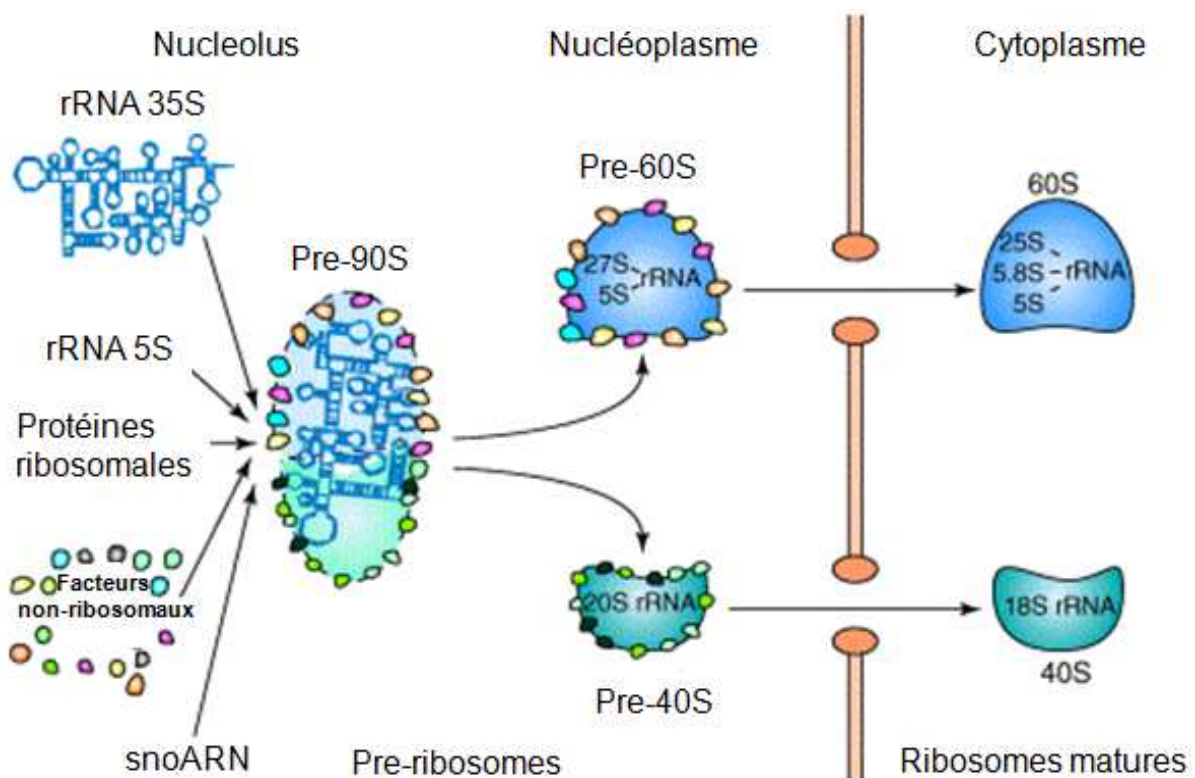


Figure 1.1 : Assemblage, maturation et export des sous-unités ribosomales 40S et 60S eucaryotiques d'après (Tschochner and Hurt, 2003).

II Le cœur universellement conservé du ribosome

La comparaison de séquences des ARNr provenant de divers organismes révèle des régions du ribosome universellement conservées. De façon marquée, ces régions comprennent les principaux domaines fonctionnels du ribosome ([Mears et al., 2002](#)). En outre, une analyse phylogénétique menée sur les ARNr a confirmé cette conservation et a permis de construire un nouvel arbre phylogénétique. Celui-ci fait ressortir les trois domaines connus du vivant. En revanche, il alimente le débat sur l'ancêtre universel commun. Cette description lui attribuerait un caractère eucaryotique alors qu'un caractère procaryotique était le plus souvent considéré ([Caetano-Anolles, 2002](#)). Une structure ribosomale minimale universellement conservée comportant les organes essentiels à la réalisation de la traduction peut donc être décrite. Les éléments de cette structure font référence à la numérotation des bases des sous-unités et à la nomenclature des protéines ribosomales d'*E. coli*.

1 La petite sous-unité

Afin de faciliter l'identification des différents domaines de la petite sous-unité ribosomale, il a été attribué une définition morphologique des domaines (**Figure 1.2**). Ainsi, on identifie une tête et un bec liés au corps, à l'épaule et à la plateforme par le cou. La petite sous-unité ribosomale contient le centre de reconnaissance de l'appariement entre le codon de l'ARNm et l'anticodon de l'ARNt. Les structures obtenues par co-cristallisation du ribosome, d'ARNt et de l'ARNm ont permis d'identifier leurs sites de liaison respectifs ([Carter et al., 2000](#); [Cate et al., 1999](#)). Il existe en réalité trois sites de liaison de l'ARNt. On compte le site A (Aminoacyl), lieu de fixation des ARNt élongateurs, le site P (Peptidyl), liant l'ARNt portant le polypeptide en croissance mais également l'ARNt initiateur, et enfin le site E (Exit) où les ARNt déacylés sont transférés avant d'être éjectés du ribosome. De plus, en calculant une carte de différence Fourier entre des structures de ribosome complexées ou non avec l'ARNm, il a pu être identifié un canal de liaison de l'ARNm. L'ARNm s'adapterait au cou de la petite sous-unité tel un collier ([Yusupova et al., 2001](#)). Les éléments à la base de ces interactions sont strictement conservés et les bases moléculaires de la liaison des partenaires ARN clairement identifiées.

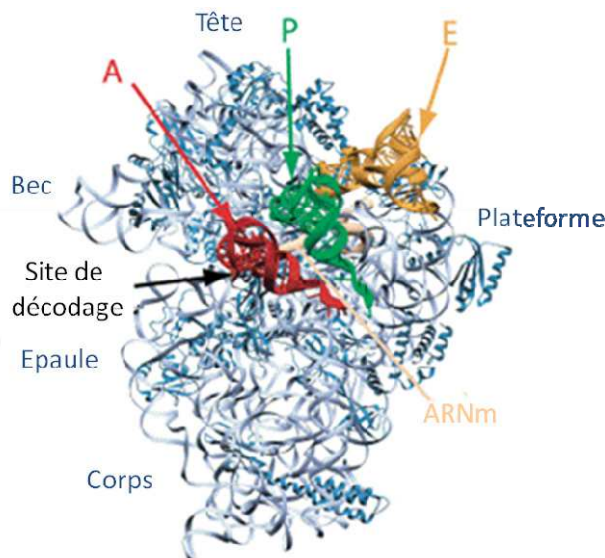


Figure 1.2 : Vue de la sous-unité 30S du ribosome bactérien depuis son interface avec la grande sous-unité d'après ([Ramakrishnan, 2002](#)). Les ARNt présents aux sites A, P et E ainsi que l'ARNm sont représentés.

a Liaison de l'ARNt initiateur

Les ARNt adoptent une structure particulière à 4 bras dite en « feuille de trèfle », maintenue par des appariements intramoléculaires entre les nucléotides le composant, conduisant à une structure tridimensionnelle en « L » (**Figure 1.3**). La première base de l'anticodon peut être modifiée : l'adénine en inosine ou l'uracile en pseudouridine. En effet, la redondance du code génétique se traduit notamment par l'usage, pour un même acide aminé, de plusieurs codons qui ne diffèrent que de leurs dernières bases. La glycine est par exemple codée par les codons GGG, GGA, GGC et GGU. Ainsi, des modifications post-transcriptionnelles de la première base de l'anticodon permettront des interactions non-canoniques (appariements « wobble » ou « bancals ») avec la dernière base du codon (base dit « wobble »).

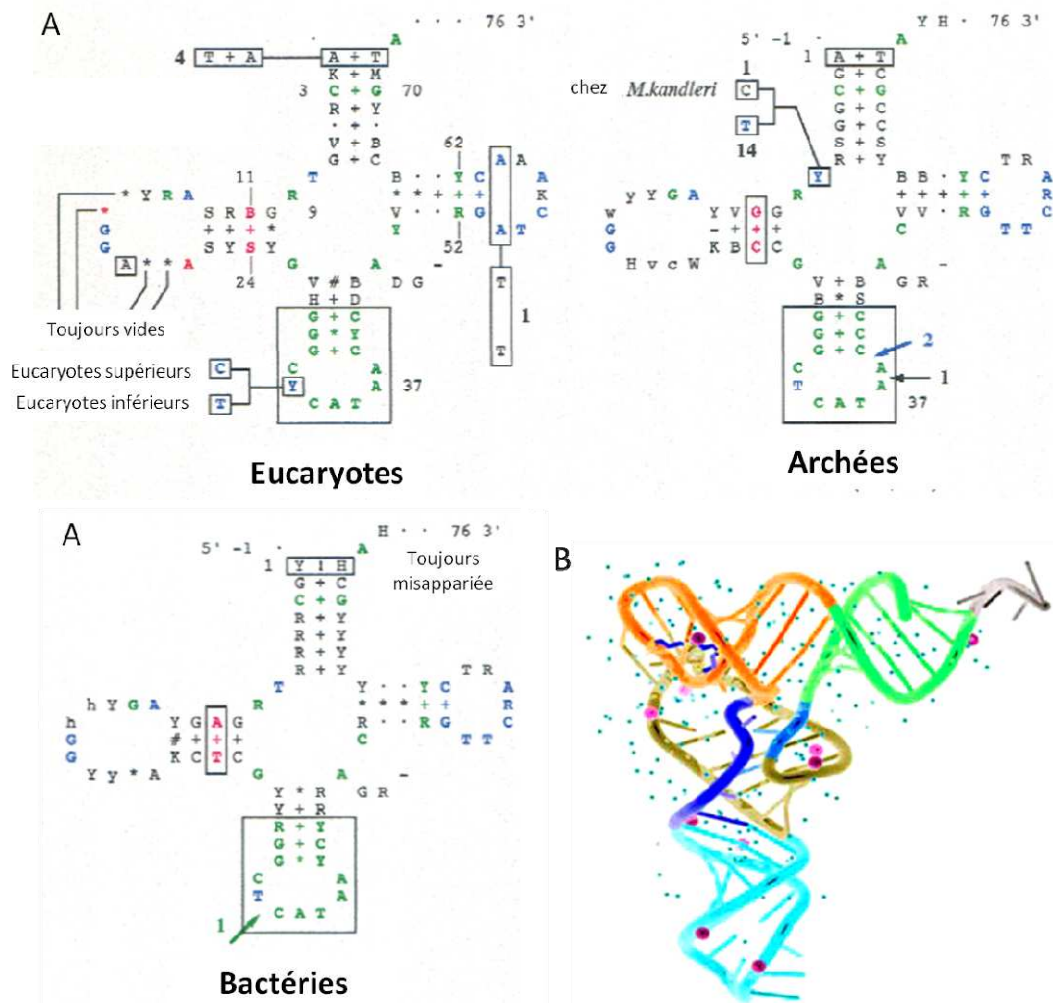


Figure 1.3 : Représentation en « feuille de trèfle » et structure tridimensionnelle de l'ARNt. (A) Représentation en « feuille de trèfle » des ARNt initiateur des Eucaryotes, des Archées et des Bactéries d'après (Marck and Grosjean, 2002). Les absences d'appariements sont indiquées par des !. Les bases colorées en bleu sont universellement conservées dans tous les ARNt, le rouge indique les bases conservées au sein d'un domaine, les bases vertes sont conservées dans tous les ARNt initiateur. Les régions encadrées correspondent aux caractères remarquables. Les encadrés à l'extérieur de la représentation indiquent des configurations alternatives et leur occurrence est notée. Les flèches indiquent la position d'introns. (B) Structure tridimensionnelle de l'ARNt^{Phe} de levure d'après (Schevitz et al., 1979). Le bras anticodon est indiqué en cyan, le bras T en orange, le bras D en jaune, la bras accepteur en vert et la tige acceptrice en gris.

Une différence notable au niveau de l'ARNt initiateur de tous les organismes réside dans la présence de 3 paires de bases G:C consécutives au sein du bras anticodon (**Figure 1.3(A)**). Ce déterminant est notamment reconnu spécifiquement au niveau du site P de la petite sous-unité ribosomale ([Seong and RajBhandary, 1987a](#)). Deux résidus universellement conservés : G1338 et A1339 assurent cette discrimination par des interactions ARN-ARN de type « A-petit sillon » ([Selmer et al., 2006](#)) (**Figure 1.4(B)**). Ces motifs d'interaction ARN-ARN ont été définis à partir de la description des interactions stabilisant le ribosome ([Nissen et al., 2001](#)). Il a pu être identifié des appariements autres que Watson-Crick ou « wobble » G:U. Quatre types de motif « A-petit sillon » sont définis en fonction des positions relatives des atomes O2' et N3 de l'adénine par rapport aux atomes O2' de la paire de base de l'hélice impliquée. La base guanine peut également satisfaire ce type de motif mais uniquement celui de type II ([Doherty et al., 2001](#)).

La base G1338 interagit avec la paire G29:C41 de l'ARNt initiateur par un motif de type II c'est-à-dire que seul l'atome N3 de la guanine est inséré dans le petit sillon alors que l'atome O2' est à l'extérieur. L'adénine 1339 réalise un motif de type I avec la paire G30:C40 au sein duquel les atomes N3 et O2' sont situés à l'intérieur du petit sillon et des liaisons non-covalentes sont créées avec les deux bases de la paire ([Berk et al., 2006](#); [Selmer et al., 2006](#)) (**Figure 1.4(B)**). Il est à noter que les bases G1338 et A1339 sont localisées au niveau de la tête de la petite sous-unité. Un mouvement rigide de rotation de la tête par rapport au cou évalué à 12° a pu être observé par comparaison des différentes structures cristallographiques publiées ([Schuwirth et al., 2005](#)). Les structures illustreraient alors différents états de stabilisation de l'ARNt au site P. En outre, les interactions impliquant le motif « A-petit sillon » stabilisent d'avantage les paires G:C par rapport aux paires C:G, A:U ou U:A ([Doherty et al., 2001](#)). Ceci confère une spécificité intrinsèque au site P pour l'ARNt initiateur, ce qui a pu être confirmé par une étude de mutagenèse des résidus G1338 et A1339 altérant la spécificité de la liaison de l'ARNt initiateur au site P ([Lancaster and Noller, 2005](#)).

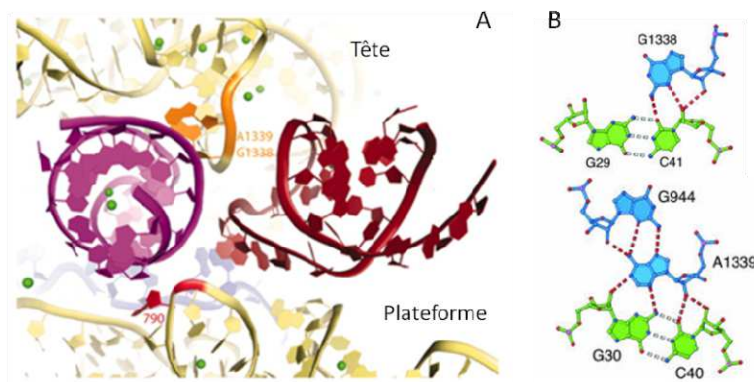


Figure 1.4 : Liaison du bras anticodon de l'ARNt au site P (A) Maintien de l'ARNt au site P (violet) entre la tête et la plateforme de la sous-unité 30S d'après ([Selmer et al., 2006](#)). L'ARNt au site E est représenté en rouge (B) Reconnaissance des paires de bases G:C par les motifs de type A-petit sillon d'après ([Berk et al., 2006](#)).

Au niveau de la plateforme, le phosphate du nucléotide A790 contacte la boucle anticodon par une liaison hydrogène avec le groupement 2'-OH de la base 38. De plus, des contacts ARN-protéines entre les protéines ribosomales S9 et des atomes d'oxygène des phosphates de l'ARNt renforcent l'interaction. L'ARNt fixé au site P est donc fermement maintenu par un réseau liant son bras anticodon avec la tête et la plateforme de la petite sous-unité (**Figure 1.4(A)**).

b Liaison de l'ARNm et contrôle de l'appariement

La petite sous-unité accommode l'ARNm par des interactions avec son squelette polyphosphaté. De façon tout à fait intéressante, l'ARNm présente un coude d'environ 45° entre les codons des sites A et P (Selmer et al., 2006; Yusupova et al., 2001). Celui-ci est maintenu par un ion magnésium coordonné par les atomes d'oxygène des groupements phosphates du troisième nucléotide du codon au site P et du premier nucléotide de celui au site A (Figure 1.5(A)). Cette liaison est renforcée par la coordination de cet ion magnésium par les nucléotides G1401 et C1402 de l'ARNr 16S. Ainsi, l'ion pontre une interaction entre des résidus de la petite sous-unité avec les nucléotides de l'ARNm situés entre le site A et le site P. Cette configuration pourrait indiquer le mode de définition du cadre de lecture et la manière dont le ribosome le maintient.

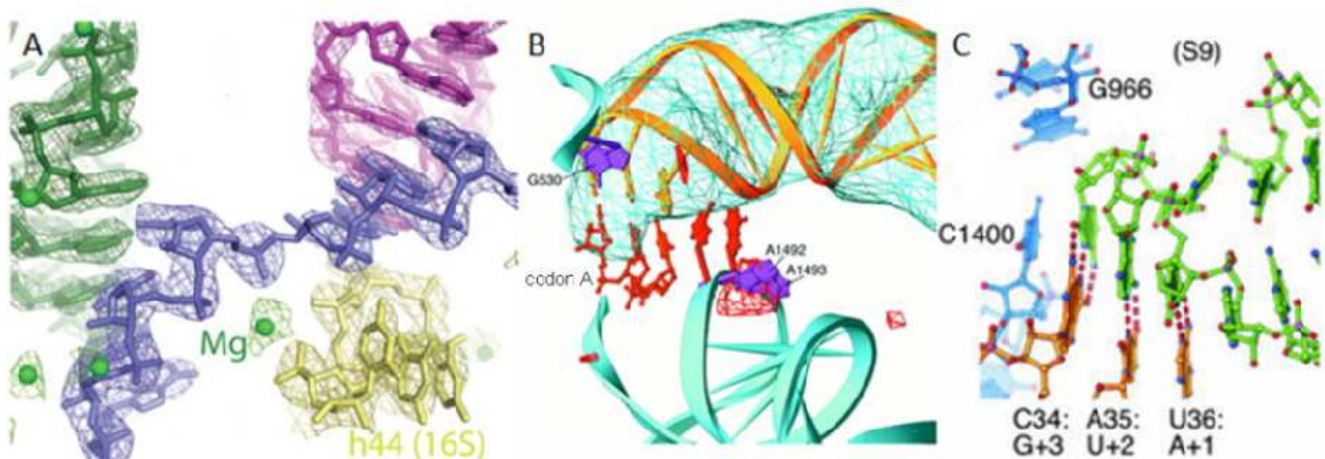


Figure 1.5 : Contrôle de l'appariement codon-anticodon aux sites A et P. (A) Vue du coude de l'ARNm (bleu) entre les sites A (dont l'ARNt est en vert) et P (dont l'ARNt est en violet) d'après (Selmer et al., 2006) (B) Reconnaitance de l'appariement entre l'ARNt au site A (doré) et le codon de l'ARNm (noté codon A). Les positions des bases A1492 et A1493 sont indiquées en présence (rouge) ou en l'absence (violet) de paromomycine qui induirait un mouvement de ces bases similaire à celui induit par la présence d'un ARNt au site A (Carter et al., 2000; Yusupov et al., 2001). (C) Reconnaitance de l'appariement codon (orange) anticodon (vert) au site P d'après (Berk et al., 2006).

Le site P

Au niveau du site P, les bases C1402, C1403 et U1498 stabilisent la liaison du squelette phosphaté du codon de l'ARNm par des liaisons hydrogènes. De plus, la première base de l'anticodon est maintenue par deux interactions de type empilement π . Ainsi, la base C1400 de la petite sous-unité ribosomale interagit avec la première base de l'anticodon et la base G966 est empilée sur le cycle ribose de ce nucléotide (Figure 1.5(C)). La troisième base du codon est donc la plus fortement surveillée.

Dans ce contexte, l'appariement le plus faible et donc la base « wobble » correspond à la première base du codon initiateur. Ceci concorde avec l'observation que le démarrage de la traduction puisse tolérer les codons GUG ou UUG (Shine and Dalgarno, 1974). En revanche, cela diffère radicalement de ce qu'impose la dégénérescence du code génétique. Le site P semble ainsi pouvoir préférentiellement sélectionner l'ARNt initiateur et stabiliser l'ARNt dans son site de liaison.

Le site A

Au niveau du site A, le mode de reconnaissance est bien différent de celui mis en place au site P. Le contrôle de l'appariement est réalisé par les résidus A1492, A1493 et G530 (**Figure 1.5(B)**). A1493 est engagé dans un motif « A-petit-sillon » de type I avec la première paire de bases codon-anticodon. L'appariement de la deuxième base est contrôlé par deux motifs de type II impliquant la base A1493 et la base du codon d'une part, et la base G530 et la base de l'anticodon d'autre part. La troisième base du codon est majoritairement maintenue par une liaison hydrogène entre son groupement 2'-OH et l'atome O6' de la base G530. Le ribose de la première base de l'anticodon est maintenu par un empilement π avec le résidu C1054 de la tête de la petite sous-unité. Ainsi, le réseau d'interaction assure une haute sélectivité vis-à-vis d'un appariement codon-anticodon de type Watson-Crick des 2 premières paires de bases alors que l'appariement au niveau de la troisième base du codon est moins surveillé.

Ces caractéristiques concordent avec la prévalence de la troisième base « wobble » pour pallier le nombre moins important d'ARNt élongateurs que de codons ([Ogle et al., 2001](#)). Le mécanisme de contrôle repose alors sur le mouvement des bases suite à la reconnaissance de l'appariement. Les bases A1492 et A1493 auparavant appariées au sein de l'hélice 44 en sont extraites pour interagir avec le petit sillon de l'hélice formée par l'appariement codon-anticodon. De même, la conformation de la base G530 est modifiée lorsqu'elle se trouve engagée dans le réseau d'interactions. Le mouvement, transmis à l'ensemble de la sous-unité, constitue une étape clé de la phase d'allongement.

Le site E

Au niveau du site E, aucun contrôle de l'appariement n'est effectué mais les ARNt déacylés sont sélectivement stabilisés et leur tige acceptrice orientée hors du centre catalytique de la grande sous-unité.

D'un point de vue dynamique, un modèle de « fermeture » de la tête de la petite sous-unité provoqué par la liaison d'un ARNt au site P et l'association des sous-unités ribosomales est de plus en plus admis. En effet, la comparaison des structures de la sous-unité 30S vide ou liée à un ARNt fait apparaître une rotation de la tête quand la sous-unité 50S se lie ([Berk et al., 2006](#); [Korostelev et al., 2006](#)). Ceci a pour effet de clore la poche du site P en rapprochant les résidus G966 et C1400 venant dès lors immobiliser l'ARNt et surtout l'ARNm dont le cadre de lecture a été choisi (**Figure 1.4(A)**). Un mouvement de ces mêmes résidus serait alors requis après la formation de la première liaison peptidique afin de réaliser la translocation de l'ARNt initiateur et de l'ARNm vers le site E.

2 La grande sous-unité

La grande sous-unité ribosomale contient le centre catalytique de l'activité peptidyl transférase. Les domaines caractéristiques de cette sous-unité sont également universellement conservés et se répartissent en deux groupes. Le premier comprend les éléments localisés au cœur de la sous-unité, à savoir le centre peptidyl transférase (PTC) et les sites de liaison des ARNt. Un deuxième groupe se compose de domaines situés à la périphérie : la protubérance centrale, la protéine L1, le bras de l'ARNr fixant L1, la protéine L11 fixant l'extension L7/L12, le bras fixant la protéine L11 et la boucle α -sarcine-ricine (α -SRL) ([Ban et al., 2000](#); [Cate et al., 1999](#)) (**Figure 1.6**).

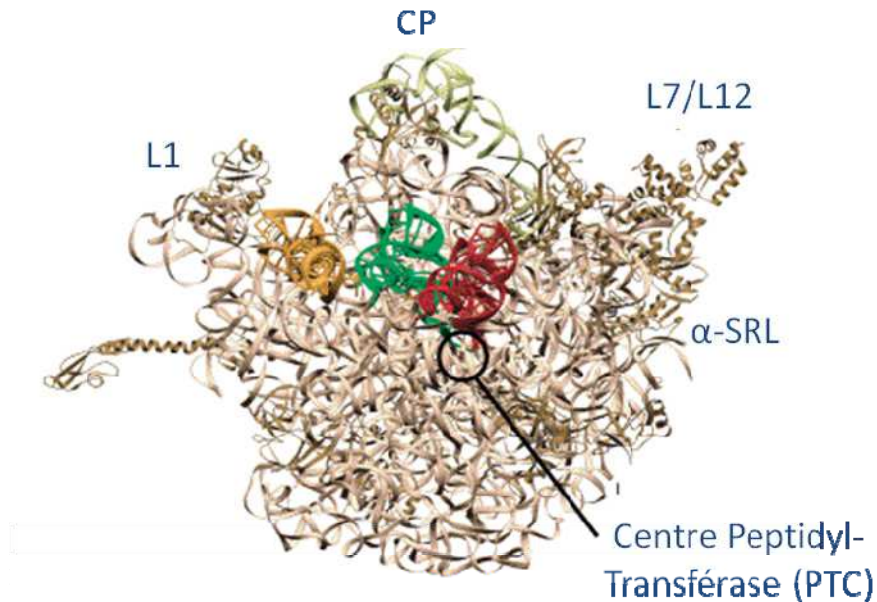


Figure 1.6 : Vue de la sous-unité 50S du ribosome bactérien depuis son interface avec la petite sous-unité d'après ([Ramakrishnan, 2002](#)). Les ARNt présents aux sites A, P et E ainsi que l'ARNm sont représentés.

a L'ARNt au site P et le centre catalytique

Le centre catalytique se situe entre les sites de liaison des ARNt aux sites A et P sur la grande sous-unité. L'extrémité 3' CCA du bras accepteur de l'ARNt fixé au site P est stabilisée par des appariements de type Watson-Crick entre les bases C74 et C75 de l'ARNt et les bases G2252 et G2251 de l'ARNr 23S, respectivement (**Figure 1.7(C)**). L'adénine 76 est empilée sur les résidus 75 et 74 et forme une interaction de type « A-petit sillon » avec la paire de bases A2450:C2063 de l'ARNr 23S. En plus de ces contacts, l'hélice 69 de l'ARNr 23S interagit avec les résidus 11 et 12 du bras D et les nucléotides 24 et 25 du bras accepteur de l'ARNt. La protéine ribosomale L5 forme des liaisons hydrogènes avec le résidu C56. Enfin, 5 ions magnésiums sont coordonnés au sein de l'ARNt et un sixième ponte les oxygènes des phosphates des nucléotides 75 et 76 avec le nucléotide 2062 de l'ARNr. Ainsi, l'ARNt lié au site P du ribosome est engagé dans de multiples contacts sur la grande sous-unité.

Le groupement 2'-OH du nucléotide A76 est suffisamment proche de la base A2451 pour participer à des liaisons hydrogènes par son atome N3 et son groupement 2'-OH. Ces interactions sont certainement à la base de l'activité peptidyl transférase. En effet, une altération de ces liaisons réduit considérablement la vitesse de formation de la liaison peptidique. Seule est requise la présence des groupements 2'-OH des résidus A76 et A2451, les bases pouvant être mutées sans grandes conséquences pour la transpeptidation ([Erlacher et al., 2005](#); [Weinger et al., 2004](#)) (**Figure 1.7(A)**). Le bras accepteur de l'ARNt présent au site A est quant à lui orienté par une interaction de type petit-sillon entre l'hélice 89 de l'ARNr 23S et l'extrémité CCA. Celle-ci est positionnée par la boucle 1492 conservée de l'ARNr 23S et l'interaction entre les bases C75 de l'ARNt et G2553 de l'ARNr 23S (**Figure 1.7(B)**).

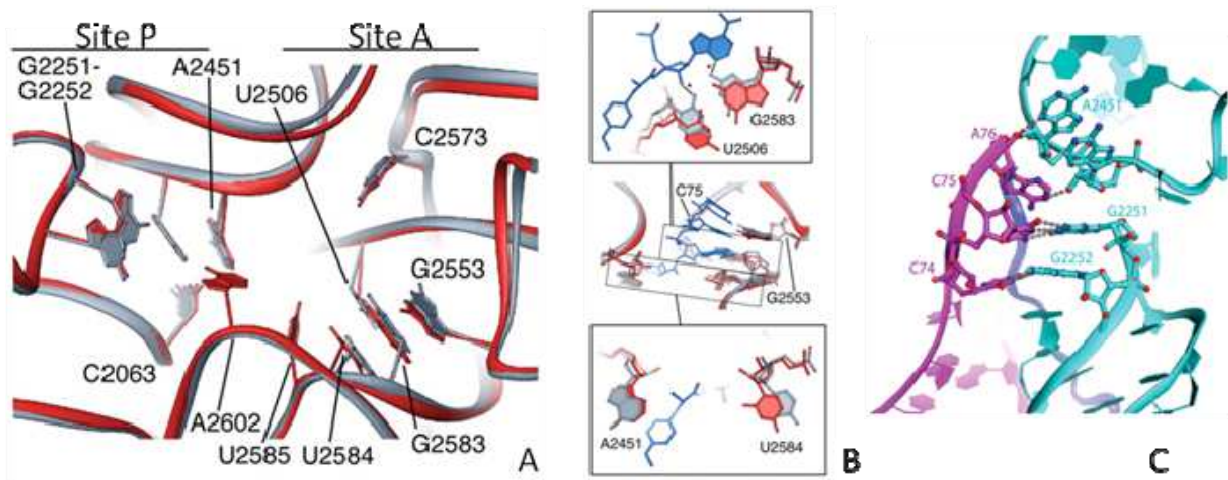


Figure 1.7 : Vue du centre peptidyl-transférase de la sous-unité 50S. (A) Superposition des sites catalytiques de la structure d'*E. coli* (gris) et *H. marismortui* (rouge). (B) Liaison de la tige CCA d'un ARNt modèle au site A (bleu) d'après ([Schuwirth et al., 2005](#)). (C) Vue de la tige acceptrice de l'ARNt lié au site P de la sous-unité 50S d'après ([Selmer et al., 2006](#)).

La chaîne polypeptidique allongée au niveau de ce centre catalytique est dirigée vers l'extérieur du ribosome par un canal composé essentiellement d'ARN. D'une longueur d'environ 100Å et d'un diamètre d'approximativement 20Å, le canal s'étend du centre catalytique jusqu'à la périphérie, en bas de la sous-unité, face à la protubérance centrale.

b La protéine L1

En périphérie de la grande sous-unité, la protéine L1 et son site de liaison sur l'ARNr interagissent avec l'ARNt lié au site E. Ainsi, un mouvement de la protéine par rapport à son site de liaison serait nécessaire pour éjecter l'ARNt déacylé, ce qui est soutenu par les différentes configurations observées dans cette région ([Yusupov et al., 2001](#)).

c Le centre d'activation des GTPases

La boucle α -SRL est le substrat des protéines toxiques α -sarcine et ricine qui modifient respectivement cette boucle en la clivant et en modifiant un nucléotide ([Endo et al., 1987](#); [Endo and Wool, 1982](#)). Les toxines altèrent ainsi la fixation et l'activation de facteurs protéiques de la traduction possédant une activité GTPase activée par le ribosome. Celui-ci ne peut alors plus assurer la traduction.

A côté de cette région, le site de liaison de la protéine L11 apparaît hautement mobile dans les différentes structures. La protéine L11 s'associe au ribosome sur un site partagé avec un groupe de protéines qui adopte une structure étendue. Il comprend une protéine L10 chez les Bactéries ancrée sur l'ARNr à côté de L11 et 4 copies de la protéine L7 ou L12 (la protéine L7 est équivalente à la protéine L12 N-acétylée) liées à L10. Des orthologues de ces protéines existent dans tous les règnes du vivant. L10 est désignée L10E chez les Archées et P0 chez les Eucaryotes et les équivalents eucaryotiques des protéines L7 et L12 appartiennent à deux familles P1 et P2. Bien que ces protéines aient divergé en séquence au cours de l'évolution, la structure de l'extension ainsi que leurs fonctions semblent préservées. En effet, ces extensions interagissent avec les domaines G des mêmes protéines qui se lient à la boucle SRL. Ainsi, ces protéines ribosomales stimuleraient le recrutement des protéines G du système de la traduction et l'activité GTPase par stabilisation de la conformation de leurs sites actifs ([Diaconu et al., 2005](#); [Savelsbergh et al., 2000](#)).

3 Liaison des sous-unités

Au cours du démarrage, l'information provenant du site de décodage doit être communiquée au site de formation de la liaison peptidique et la structure du ribosome doit être suffisamment mobile pour permettre la translocation. La communication entre les sous-unités peut être assurée par les ligands du ribosome présents simultanément sur les deux sous-unités à savoir les ARNt ou bien par les liens qui connectent les sous-unités. Notamment, on observe que les résidus impliqués dans la formation de ponts entre les sous-unités sont significativement conservés dans les trois domaines du vivant. De plus, les ions magnésiums jouent un rôle crucial dans la formation de ponts entre les sous-unités. En comparant les structures cristallographiques de sous-unités 50S bactériennes et archéennes, il s'avère que près de 40% des sites de liaison des ions magnésiums sont conservés dans tous les domaines du vivant et après l'association des sous-unités ribosomales ([Klein et al., 2004](#); [Schuwirth et al., 2005](#)). L'obtention de structures de ribosomes à des résolutions de plus en plus élevées a permis de décrire précisément l'ensemble des ponts entre les sous-unités ([Selmer et al., 2006](#); [Spahn et al., 2001](#); [Yusupov et al., 2001](#)). Ces études ont fait apparaître une diversité dans la nature des interactions mises en jeu. Ces liaisons plus ou moins fortes concordent avec l'idée selon laquelle certains de ces ponts seraient amenés à se rompre et se former à nouveau au cours de la traduction.

Ainsi, 12 réseaux d'interactions comportant 30 contacts liant les sous-unités ont pu être identifiés dans toutes les structures (**Figure 1.8**). La plupart de ces contacts repose sur des interactions ARN-ARN localisées au niveau de la plateforme et de l'avant-dernière hélice de la petite sous-unité (hélice 44). De même, les partenaires de ces interactions tout ARN sont également concentrés au niveau du centre de la grande sous-unité. Ces réseaux reposent majoritairement sur des interactions de type petit-sillon/petit-sillon mais aussi des contacts impliquant grand-sillons, boucles et squelettes polyphosphatés. Néanmoins, certains réseaux sont renforcés par des interactions de type ARN-protéine. D'autres ponts ne sont assurés que par ce type d'interaction et l'un d'entre eux implique uniquement une interaction entre protéines.

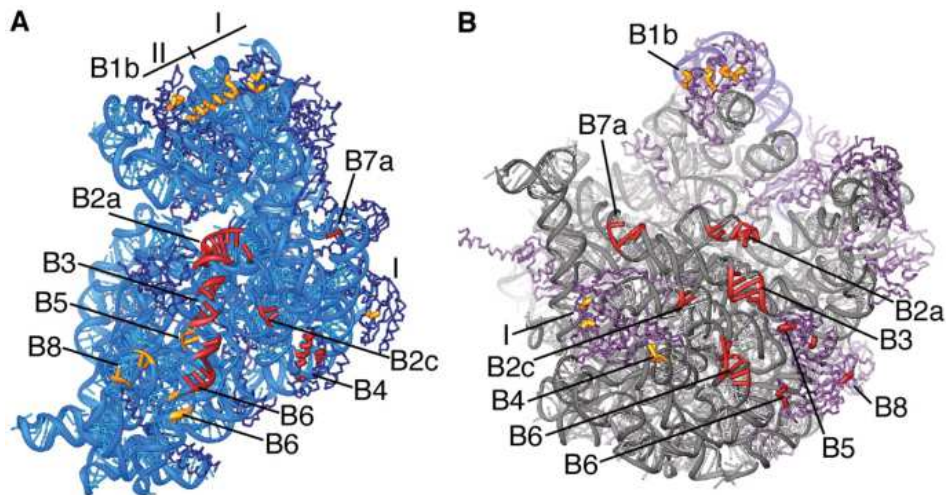


Figure 1.8 : Vue des ponts entre les sous-unités 30S et 50S d'après (Schuwirth et al., 2005). (A) Contacts à l'interface de la petite sous-unité colorés suivant le type d'interaction (doré : liaison à une protéine ; rouge : liaison à l'ARNr). L'ARNr et les protéines ribosomales sont respectivement colorées en bleu clair et bleu foncé. Les contacts I et II sont exclusifs et correspondent à deux orientations différentes de la tête observée au sein des ribosomes 70S. (B) Contacts à l'interface de la grande sous-unité colorés comme précédemment. L'ARNr 23S et les protéines ribosomales sont respectivement colorés en gris et magenta. L'ARNr 5S apparaît en violet.

a Liaison des sous-unités : la tête, un élément mobile

Deux ponts permettent de connecter la tête de la petite sous-unité avec la protubérance centrale de la grande sous-unité. Le premier auquel a été donné le nom de « doigt du site A » (« A-site finger ») est maintenu par l'interaction entre le résidu 890 de la boucle apicale d'une longue hélice de la grande sous-unité (hélice 38 de l'ARNr 23S) et une séquence basique conservée de la protéine ribosomale S13 (Figure 1.8, Pont B1b). L'hélice 38 s'étend sur la grande sous-unité de la protubérance centrale jusqu'au niveau du milieu de la tête. Comme nous l'avons vu, la tête est un élément mobile du ribosome probablement impliqué dans la translocation. Cette hélice pourrait être un médiateur important pour transmettre ou induire un mouvement de la tête de la petite sous-unité. La protéine S13 participe également à un deuxième pont, le seul impliquant deux protéines (pont B1a n'apparaissant pas sur la Figure 1.8). Elle se lie par sa queue N-terminale à la protéine L5 de la grande sous-unité, impliquée dans l'accommodation de l'ARNt au site P. Cette interaction pourrait aussi avoir un rôle dans l'échange de l'information au sein du ribosome. En effet, différentes conformations de ce pont ont pu être observées à différents stades de l'allongement (Frank et al., 2007).

b Liaison des sous-unités : des réseaux d'interactions à la fonction

Les autres ponts formés entre le corps et la plateforme de la petite sous-unité et la grande sous-unité se répartissent dans trois directions regroupant respectivement les ponts B2a/B3/B5/B6, B2c/B4 et B7a/I de la Figure 1.8.

Les ponts B2a/B3/B5 forment le réseau le plus proche de l'épaule qui suit la direction de l'hélice 44. Cette hélice est impliquée dans la formation de 4 ponts avec 3 hélices de l'ARN 23S et la protéine ribosomale L14 de la grande sous-unité. L'hélice 44 s'étend sur toute l'interface avec la grande sous-unité et contient des résidus que l'on a vu fortement impliqués dans les liaisons des ARNt aux sites A et P et dans le décodage (C1400-C1403 avec l'ARNt du site P, A1492-A1493 avec l'ARNt du site A). De plus, une de ces interactions lie l'hélice 44 de l'ARNr 16S à l'hélice 69 de l'ARNr 23S de la grande sous-unité impliquée dans la stabilisation de l'ARNt au site P sur la sous-unité 50S (Figure 1.8, Pont B2a). Or, la résolution de structures par cryomicroscopie électronique de ribosomes à divers stades

d'allongement soutient un modèle d'avancement en cliquet ([Frank and Agrawal, 2000](#); [Frank et al., 2007](#)). La translocation serait réalisée en deux étapes. Dans un premier temps, la translocase EF-G bactérienne liée au GTP se fixe sur le ribosome. Cela induit une rotation de la sous-unité 30S par rapport à la grande sous-unité, ouvrant le canal de liaison de l'ARNm et positionnant les ARNt aux sites A et P dans une conformation hybride. Leurs boucles anticodons sont translatées aux sites P et E respectivement sur la petite sous-unité 30S alors qu'ils restent aux sites A et P de la sous-unité 50S. L'hydrolyse du GTP porté par EF-G provoque alors la translocation des ARNt au regard de la sous-unité 50S. Enfin, l'éjection de la translocase du complexe induit une rotation inverse qui restaure les ponts entre les sous-unités, permettant la liaison d'un nouvel ARNt au site A. La structure obtenue par cryomicroscopie électronique d'un ribosome eucaryotique avec la translocase eEF2 confirme l'universalité du processus ([Spahn et al., 2004a](#)). Ces mouvements peuvent certainement être induits intrinsèquement par le ribosome. Par ses multiples interactions et sa position, l'hélice 44 de la petite sous-unité ribosomale pourrait avoir un rôle clé en couplant le contrôle d'un appariement correct codon-anticodon à une stimulation du mouvement relatif des sous-unités. Toutefois, les facteurs d'allongement permettraient de stimuler les différentes étapes.

Trois ponts forment ensuite le réseau central. Ceux-ci impliquent des hélices de l'ARNr 23S de la grande sous-unité et des partenaires ARN et protéines de la petite sous-unité (incluant les ponts B2c et B4, le troisième pont B2b n'apparaît pas dans la structure présentée dans la **Figure 1.8**).

La troisième ligne, la plus proche du site de liaison de la protéine L1, et donc au niveau de la plateforme, comprend une interaction ARN-ARN (pont B7a) et une interaction ARN-protéine (pont I). La protéine L2 de la grande sous-unité impliquée dans ce contact est cruciale pour l'association des sous-unités ribosomales et pourrait être importante pour l'activité peptidyl transferase ([Diedrich et al., 2000](#)).

La structure ribosomale ainsi décrite possède tous les éléments nécessaires à la traduction, à savoir : des sites de liaison de l'ARNm, de l'ARNt, un contrôle de l'appariement codon-anticodon, une mobilité intrinsèque et un centre catalytique de formation de liaison peptidique. Alors que les ARNm possèdent une séquence signal à leurs extrémités 5' pour stimuler leur recrutement au ribosome, une classe d'ARNm sans séquence promotrice (« leaderless mRNA ») a pu être identifiée. Dans ce cas, le codon de démarrage est placé en 5' ou après quelques nucléotides. Ces ARNm, majoritairement présents chez les Bactéries, les Archées et les mitochondries de Mammifères, sont efficacement traduits *in vivo*. Il a pu être montré dans le système bactérien que la traduction *in vitro* de ces ARNm peut démarrer en présence des seuls ribosome et ARNt initiateur méthionylé ([Moll et al., 2002](#); [Moll et al., 2004](#); [O'Donnell and Janssen, 2002](#)). En fait, ces ARNm peuvent s'associer à un ribosome 70S. L'ARNt initiateur méthionylé stimule cette liaison ([Udagawa et al., 2004](#)). De plus, un ARNm bactérien sans séquence promotrice peut être traduit par un ribosome 80S de Mammifère en présence d'un ARNt initiateur méthionylé ([Andreev et al., 2006](#)). Ces éléments pourraient fournir des pistes sur l'existence d'un système primitif comportant un nombre réduit de protéines. Seuls les deux sous-unités ribosomales, l'ARNm et l'ARNt initiateur auraient permis d'assurer le démarrage de la traduction.

Il faut toutefois souligner que la reconstitution *in vitro* de systèmes de traduction révèle que la traduction des ARNm sans séquence promotrice est stimulée par les facteurs de démarrage et notamment par IF2, le facteur impliqué dans le recrutement de l'ARNt initiateur au ribosome. L'apparition de ces facteurs aurait ainsi permis d'améliorer l'efficacité de la traduction. La description d'éléments du ribosome apparus au cours de l'évolution nous donnera les premiers éléments justifiant cette idée.

III Différences entre ribosomes de diverses origines

D'un point de vue général, les régions les moins conservées du ribosome sont trouvées à la périphérie des sous-unités. Celles-ci seraient apparues afin de satisfaire les contraintes de l'environnement sur les différents organismes. Les ribosomes doivent en effet réaliser la traduction avec les meilleures fidélité et efficacité possibles. En outre, leur action est couplée à des fonctions spécifiques des organismes, intervenant en aval de la traduction (transport des protéines, repliement des protéines, maturation N-terminale des protéines). Les nucléotides supplémentaires de l'ARNr des Eucaryotes se répartissent au niveau de régions de longues insertions dits segments d'extension qui ne perturbent pas la structure du cœur fonctionnel. Les différentes structures de ribosomes 80S de levure et humains obtenues par cryomicroscopie électronique ont permis de clairement identifier ces segments ([Schuler et al., 2006](#); [Spahn et al., 2001](#); [Spahn et al., 2004b](#)). Il est à noter que les différences observées entre ces structures eucaryotiques sont toutes localisées au niveau de quelques segments d'extension.

En outre, les études phylogénétiques portant sur les protéines ribosomales révèlent une variabilité importante au sein des trois domaines du vivant. En comparant 102 familles de protéines ribosomales issues de 66 espèces, Lecompte et al. ont montré que seules 34 familles sont universellement conservées. 33 familles de protéines sont communes aux Eucaryotes et aux Archées et absentes chez les Bactéries. Une famille de protéines est spécifique des Archées alors que 11 familles existent uniquement chez les Eucaryotes. Enfin, 23 familles de protéines ribosomales sont exclusivement présentes chez les Bactéries. L'analyse génétique poussée conduit à des résultats difficiles à interpréter. D'une part, les Archées semblent posséder une version simplifiée du jeu de protéines ribosomales des Eucaryotes. Les gènes des protéines manquantes auraient été perdus au cours de l'évolution et l'effet aurait été contrebalancé par un enrichissement de l'ARNr. Ceci s'inscrirait dans un contexte où l'ancêtre universel aurait une origine eucaryotique. En revanche, la franche délimitation entre le jeu bactérien et les jeux eucaryotique et archéen, associée à un jeu de protéines ribosomales universellement conservées, serait plus en faveur de l'hypothèse d'un ancêtre d'origine procaryotique ([Lecompte et al., 2002](#)).

1 Les extensions de la petite sous-unité

Au sein de la petite sous-unité eucaryotique, une nette subdivision apparaît au niveau de la partie inférieure du corps (**Figure 1.9**). Le parallèle morphologique vaut à ces deux extensions les dénominations de « pied droit » et « pied gauche ». La majorité de la densité supplémentaire est concentrée sous la plateforme, comprenant le « pied gauche » et un « lobe postérieur ». Enfin, sont observées au niveau de la tête une expansion du bec ainsi qu'une protubérance orientée vers le solvant formant le « lobe de la tête ».

a Les extensions de l'ARNr

Des différences notables sont apparues en inspectant l'épaule de la petite sous-unité du ribosome. Chez les Bactéries, l'épaule est formée par les hélices 16 et 17 de l'ARNr 16S. L'hélice 16 est repliée sur l'hélice 18 et l'hélice 17 occupe sur la petite sous-unité une région qui s'étend depuis l'épaule jusqu'en bas du corps. En revanche, au sein de l'épaule des ribosomes eucaryotiques, l'extrémité de l'hélice 16 est orientée vers la partie exposée au solvant. L'hélice 17, plus courte dans ces organismes, est fortement coudée et se replie sur la boucle terminale de l'hélice 21. Ces modifications intervenant au niveau de l'épaule sont certainement impliquées dans le fonctionnement du ribosome et pourraient intervenir dans le mécanisme de fermeture de la tête vers l'épaule.

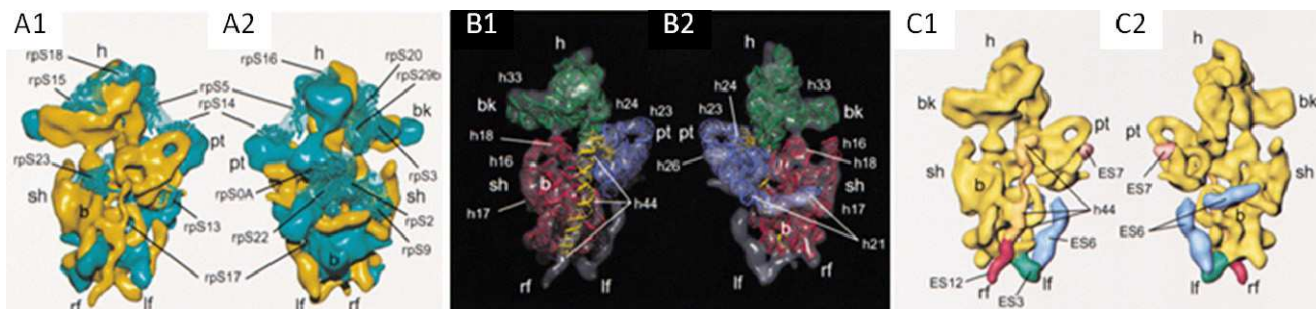


Figure 1.9 : Partition ARN/Protéines ribosomales de la sous-unité 40S de *S. cerevisiae* d'après (Spahn et al., 2001). Les vues 1 et 2 correspondent respectivement à l'interface entre les sous-unités et la partie accessible au solvant (h : tête, sh : épaule, pt : plateforme, b : corps, lf : pied gauche ; rf : pied droit). (A) La partition ARN est colorée en jaune, les protéines en bleu, les protéines identifiées par homologie en bleu transparent. (B) Superposition de la densité en ARNr montrée en transparence avec la structure modèle de l'ARNr 18S universellement conservé. (C) Visualisation des segments d'expansion par rapport à l'ARNr 18S conservé montré en jaune.

L'hélice 44, élément central du processus de la traduction, présente une insertion qui l'étend vers le bas de la sous-unité. L'hélice 44 présente alors un coude qui apparaît comme une protubérance orientée vers l'épaule et exposée au solvant. Cette dernière crée le « pied droit » (**Figure 1.9, ES12**). De plus, l'hélice 21, la plus longue insertion chez les Eucaryotes, prend naissance au centre de la plateforme d'où elle se subdivise en deux hélices de tailles dissemblables. L'hélice la plus courte crée un « lobe postérieur » qui s'étend vers l'épaule (**Figure 1.9, ES6**). La deuxième extension suit la direction du bas de la plateforme et l'excroissance forme le « pied gauche » (**Figure 1.9, ES3**). L'extrémité de cette hélice interagit avec une autre insertion située dans la région du « pied gauche ». Cette dernière insertion est encore plus importante dans le cas du ribosome humain et représente plus de la moitié des résidus additionnels de cette sous-unité comparée à l'homologue de levure. Ainsi, au sein des ribosomes eucaryotiques, la plateforme contacte la partie basse du corps de la petite sous-unité par l'intermédiaire de contacts ARN-ARN additionnels. Cette configuration est totalement spécifique de ces organismes.

b Les protéines ribosomales

Chez les Eucaryotes, la forme du bec de la petite sous-unité est également très affectée par rapport à sa configuration chez les Procaryotes. Des amas de densité électronique attribués à des protéines ribosomales se substituent à l'ARNr présent chez les Bactéries et augmentent la taille du bec par rapport au ribosome procaryotique. De même, une extension également identifiée comme un composant protéique forme le « lobe de la tête ». De plus, la densité liant la plateforme à la tête se trouve fortement élargie lorsque le ribosome eucaryotique est comparé au ribosome bactérien. Il en est déduit un contact plus important entre les protéines qui composent cette interaction. Fonctionnellement, cette densité fait partie des éléments qui forment le canal de sortie de l'ARNm. Au niveau de l'embranchement de l'hélice 21, une expansion de la densité électronique de la protéine S2 spécifique aux Eucaryotes est observée. Sous l'hélice 21, un amas de densité se substitue à celle de la protéine S16 bactérienne. Cette occupation supplémentaire pourrait expliquer les changements de l'ARNr importants observés au niveau de l'épaule. Les changements de conformation des hélices 16 et 17 pourraient résulter de l'apparition de ce cluster de protéines ou bien ce dernier serait apparu suite à la modification de l'ARNr. Enfin, deux groupes de protéines sont présents en contrebas des « pieds » de la sous-unité. Celui situé au niveau du « pied gauche » occupe en fait l'espace où est liée la protéine S20 bactérienne qui n'a pas d'équivalent chez les Eucaryotes. Ainsi, le défaut d'homologues de protéines ribosomales bactériennes chez les Eucaryotes semble être contourné par l'intervention de nouveaux partenaires occupant des sites similaires. Néanmoins, deux protéines bactériennes demeurent « orphelines » (S6 et S18). Localisées en périphérie au niveau de

la plateforme du ribosome bactérien, l'absence de ces protéines accessibles au solvant pourrait être importante fonctionnellement. On peut notamment imaginer que ces sites de liaison soient utilisés par d'autres acteurs chez les Eucaryotes.

2 Les extensions de la grande sous-unité

L'ajustement de la structure de la sous-unité 50S archéenne dans la densité électronique de la sous-unité eucaryotique a nécessité de repositionner certains éléments. On compte notamment des hélices au niveau du site E, la boucle sarcine-ricine et l'hélice sur laquelle est fixée la protéine L11 (sur laquelle est fixé le groupe de protéines P0:P1:P2). Ces changements pourraient révéler des caractères spécifiques aux ribosomes eucaryotiques mais plus probablement représenter des changements de conformations liés à l'association des sous-unités ribosomales ou encore indiquer différents états fonctionnels du ribosome 80S. Les différentes configurations observées dans les régions des protéines L1 et L11 concordent avec la mobilité requise pour le recrutement des GTPase et l'éjection de l'ARNt au site E. En revanche, l'hélice 58 de l'ARNr 25S (ou 23S) semble occuper une position différente en fonction des espèces. Par exemple, chez les Bactéries *E. coli* et *Thermus thermophilus*, cette hélice prend naissance au niveau du corps et remonte en suivant la protéine L2. Par contre, chez l'Archée *Haloarcula marismortui* et chez les Eucaryotes *Saccharomyces cerevisiae* et *Homo sapiens*, cette hélice reste localisée sous le site d'ancrage de la protéine L1. Notamment, la boucle apicale de cette hélice se voit déplacée de plus de 50Å en comparant les structures. Les implications fonctionnelles de ces configurations restent à élucider. Globalement, les insertions au niveau de l'ARNr affectent deux zones : la première est positionnée derrière la protubérance centrale et le site d'ancrage de la protéine L11 et une seconde s'étend du bas de la sous-unité jusqu'à la protubérance de la protéine L1 (**Figure 1.10**). De plus, ces extensions se présentent moins sous la forme de protubérances que de structures impliquées dans des interactions maintenant la structure du ribosome. Deux de ces segments sont impliqués dans la formation de ponts entre les sous-unités.

a Les extensions de l'ARNr

Au niveau du domaine de liaison de la protéine L11, l'hélice 25 est considérablement allongée et se trouve fusionnée à une autre extension de l'ARNr. De même, une insertion au niveau de l'hélice 98 interagit avec cette même extension et l'interaction se prolonge jusqu'à l'hélice 25. Dans la structure du ribosome humain, cette région se trouve fortement élargie et forme des extensions accessibles au solvant (**Figure 1.10, ES7**). En contrebas de la protubérance centrale, une petite insertion au niveau de l'hélice 38 a pu être résolue. Cette modification doit avoir une signification mécanistique puisque nous avons vu que par sa position et son réseau d'interactions, cette hélice pourrait constituer une des bases du transfert de l'information entre les sous-unités.

La deuxième région modifiée, sous le site de fixation de la protéine L1, présente 6 extensions d'ARN. Celles-ci sont majoritairement impliquées dans des contacts additionnels avec d'autres domaines de la grande sous-unité indiquant une stabilisation accrue (**Figure 1.10, ES7/11/15**). Un amas de densité contacte l'hélice qui porte la protéine L1. Ce contact pourrait être important pour le contrôle du mouvement requis de la protéine lors du cycle d'allongement. De plus, ce même segment est impliqué dans la formation d'une interaction avec la petite sous-unité et contacte d'autres domaines de la grande sous-unité. Tout ceci renforce l'importance de cette modification bien que son implication reste encore à démontrer.

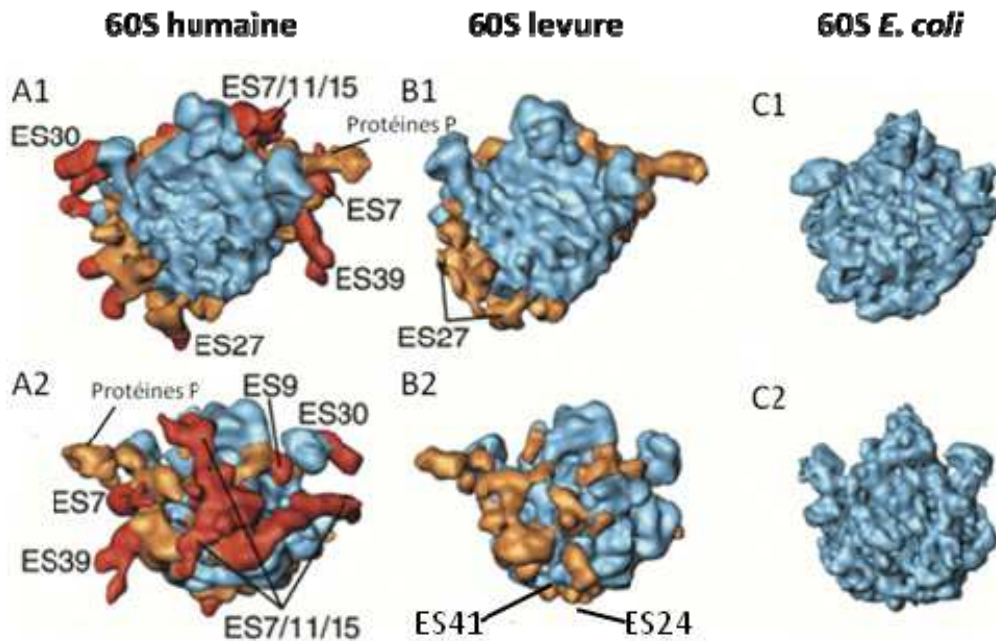


Figure 1.10 : Vues des grandes sous-unités ribosomales de différents organismes d'après ([Spahn et al., 2004b](#)). Les vues 1 et 2 se réfèrent respectivement à l'interface avec la petite sous-unité et la partie accessible au solvant. (A), (B), (C) Le cœur universellement conservé est coloré en bleu, les extensions du ribosome de levure identifiées en orange et celles caractéristiques du ribosome humain représentées en rouge.

Une des plus longues insertions au sein des ribosomes eucaryotiques se situe au niveau de l'hélice 63. Celle-ci s'étend depuis le bas de la sous-unité du côté de la protubérance P0:P1:P2 et la traverse de part en part pour terminer sa course sous le site de fixation de la protéine L1 (**Figure 1.10, ES27**). Cette extension adopterait une configuration hélicale chez la levure. Néanmoins, la grande sous-unité du ribosome humain présente une insertion encore plus importante mais mal résolue ([Morgan et al., 2000](#)). L'hypothèse la plus probable serait une grande mobilité de cette région. L'extension pourrait adopter différentes configurations, limitant alors sa visualisation ([Spahn et al., 2004b](#)). Dans une structure où ce fragment est identifiable, deux orientations possibles sont proposées. La première place l'extension proche du canal de sortie de la chaîne polypeptidique alors que la seconde la situe près du site de liaison de la protéine L1 et donc du site E. Ainsi, cette extension pourrait coordonner l'intervention de facteurs impliqués dans des événements post-traductionnels telles les protéines chaperonnes ou encore les enzymes de modification ([Beckmann et al., 2001](#); [Morgan et al., 2002](#)).

La structure du ribosome 80S de levure a mis en évidence la présence de deux autres extensions en contrebas de la sous-unité (**Figure 1.10, ES24 et ES41**). Celles-ci forment deux « doigts », l'un impliqué dans un contact nouveau entre sous-unités alors que le second permet l'accrochage d'un partenaire protéique ([Beckmann et al., 2001](#)).

Les différences les plus notables entre les ARNr des ribosomes humains et de levure se concentrent dans la région derrière la protubérance centrale et la protubérance P0:P1:P2. Un très grand amas de densité électronique correspondant à de l'ARN en forme de V n'a pas d'équivalent chez la levure. Le rôle de cette extension demeure énigmatique.

b Les protéines ribosomales

Les protéines ribosomales homologues aux protéines archéennes et bactériennes ont pu être aisément placées dans la densité électronique obtenue par cryomicroscopie électronique des ribosomes Eucaryotes. Les protubérances P0:P1:P2 et L10:L7:L12 adoptent différentes conformations. Le groupe de protéines eucaryotiques n'a pas une structure aussi étendue que dans le cas bactérien mais il présente un lobe latéral près de la liaison de la protéine P0 à la protéine L11. En revanche, l'attribution de la densité électronique aux protéines qui n'ont pas d'équivalent dans les autres domaines du vivant est très limitée. En effet, nous disposons de très peu d'informations concernant la structure de ces protéines. Il peut néanmoins être constaté que ces amas de densités semblent pouvoir entrer en contact avec plusieurs partenaires protéiques ou ribosomaux. Il peut alors être envisagé que ces protéines aident essentiellement à stabiliser la structure du ribosome ou bien alors permettent l'interaction avec d'autres partenaires de l'environnement cellulaire eucaryotique.

3 Les ponts entre les sous-unités

L'examen des structures de ribosomes 80S a permis d'identifier la présence de 4 ponts supplémentaires contribuant à maintenir les sous-unités ensemble (**Figure 1.11**). Les deux premiers ponts, eB10 et eB11, viennent en fait prolonger la ligne décrite par les ponts engageant l'hélice 44 (ligne la plus proche de l'épaule de la petite sous-unité). Ces derniers reposent majoritairement sur des contacts ARN-ARN mais l'analyse de la densité électronique indique que les ponts pourraient être également maintenus par des protéines. L'identification des partenaires protéiques devrait permettre d'expliquer plus précisément pourquoi ces contacts sont spécifiques aux Eucaryotes.

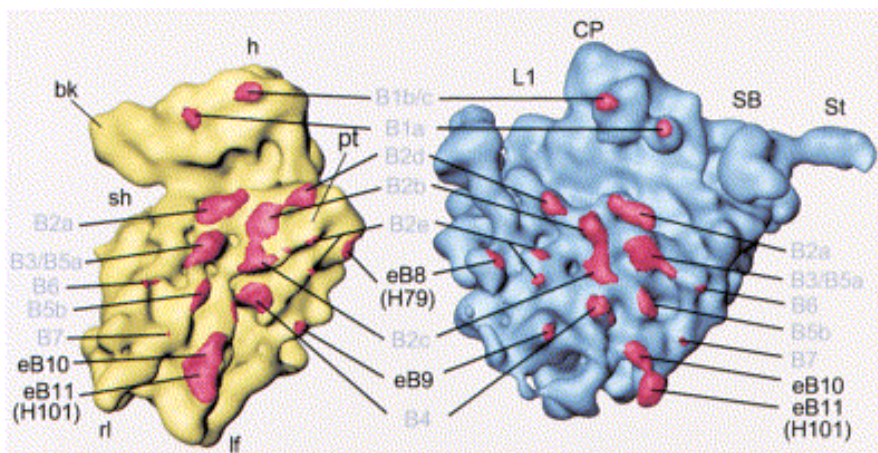


Figure 1.11 : Ponts entre les sous-unités 30S (jaune) et 50S (bleu) vues depuis leurs domaines de liaison d'après ([Spahn et al., 2001](#)) (h : tête, sh : épaule, pt : plateforme, lf : pied gauche ; rf : pied droit ; SB : point d'ancrage des protéines P notées St). Les ponts caractéristiques des Eucaryotes sont notés eB.

Le réseau de ponts situé sur la ligne jouxtant la protéine L1 présente une interaction supplémentaire, eB9, faisant intervenir des composants uniques aux systèmes eucaryotiques. La nature exacte de ce pont n'a pu être clairement identifiée mais une protéine inconnue interagirait avec un partenaire ARN ou une protéine ; la longue insertion au niveau de l'hélice 21 pourrait être impliquée.

Enfin, en contrebas de la protubérance formée par la protéine L1, une protéine ribosomale de la petite sous-unité participe à une liaison avec une extension de l'ARNr de la grande sous-unité formant le pont eB8. Comme il a déjà été précisé, ces contacts jouent certainement un rôle dans le contrôle de la mobilité de la protéine L1 et donc dans la capacité du ribosome à éjecter l'ARNt lié au site E.

PARTIE B LE DEMARRAGE DE LA TRADUCTION

La description des structures ribosomales des différents organismes nous a permis de percevoir la complexité de l'organisation fonctionnelle de ce ribozyme. L'évolution a cependant fait apparaître des facteurs protéiques afin d'aider le ribosome à rendre le processus le plus efficace et le plus fidèle possible. De plus, ces facteurs interviennent également dans le contrôle de son activité. La comparaison des jeux de facteurs de démarrage indique des différences notables entre les Bactéries, les Archées et les Eucaryotes (**Tableau 1**). Nous allons dans un premier temps décrire les systèmes de démarrage dans les trois domaines du vivant en ne donnant qu'un aperçu des étapes clés du système eucaryotique qui sera plus étoffé dans la seconde partie.

Facteurs eucaryotiques	Facteurs archéens	Facteurs bactériens	Facteurs universellement conservés
<u>eIF1</u>	aIF1	<u>IF3</u>	e/aIF1/IF3C
<u>eIF1A</u>	<u>aIF1A</u>	<u>IF1</u>	e/aIF1A/IF1
<u>eIF2</u>	<u>aIF2</u>		
<u>eIF2B</u>	<u>aIF2B</u>		
<u>eIF3</u>			
<u>eIF4A</u>			
<u>eIF4B</u>			
<u>eIF4E</u>			
<u>eIF4G</u>			
eIF4H			
<u>eIF5</u>			
eIF5B	<u>aIF5B</u>	<u>IF2</u>	e/aIF5B/IF2
<u>eIF6</u>			
<u>PABP</u>			

Tableau 1.1 : Facteurs du démarrage de la traduction dans les trois domaines du vivant. Les facteurs dont la structure a été résolue sont soulignés en traits pleins, les traits pointillés correspondent à des structures de domaines ou sous-unités résolues.

I Le démarrage bactérien

Le ribosome bactérien est assisté par trois facteurs de démarrage : IF1, IF2 et IF3. Le démarrage de la traduction a lieu sur la petite sous-unité seule (**Figure 1.12**). Il doit donc être ajouté une étape préalable de dissociation des sous-unités ribosomales. Dans les conditions physiologiques, l'équilibre existant entre sous-unités associées et sous-unités isolées favorise en effet la formation du ribosome 70S. L'intervention de partenaires protéiques assure alors le déplacement de l'équilibre. Ceux-ci agissent en se fixant sur la petite sous-unité au niveau de l'interface avec la grande sous-unité. IF3 possède une telle activité, stimulée par IF1 ([Grunberg-Manago et al., 1975](#)). En outre, des études portant sur l'étape de terminaison indiquent que le facteur IF3 intervient à la fin de la traduction, dans le processus de recyclage du ribosome, avec le facteur de recyclage RRF (Ribosome Recycling Factor) et la translocase EF-G (Elongation Factor G). Suivant l'hydrolyse du GTP porté par le facteur EF-G, une sous-unité 50S libre et une sous-unité 30S comportant le dernier ARNt déacylé présent au site P sont ainsi créées. Le facteur IF3 stimule alors l'éjection de l'ARNt déacylé, générant une petite sous-unité sans substrat ([Agrawal et al., 2004](#); [Karimi et al., 1999](#)).

La petite sous-unité ainsi libérée peut accueillir l'ARNm, l'ARNt initiateur méthioninylé et formylé (fMet-ARNt^{fMet}) et les facteurs de démarrage. La chronologie des étapes de liaison des partenaires sur la sous-unité 30S reste floue même si les mécanismes moléculaires individuels sont bien caractérisés.

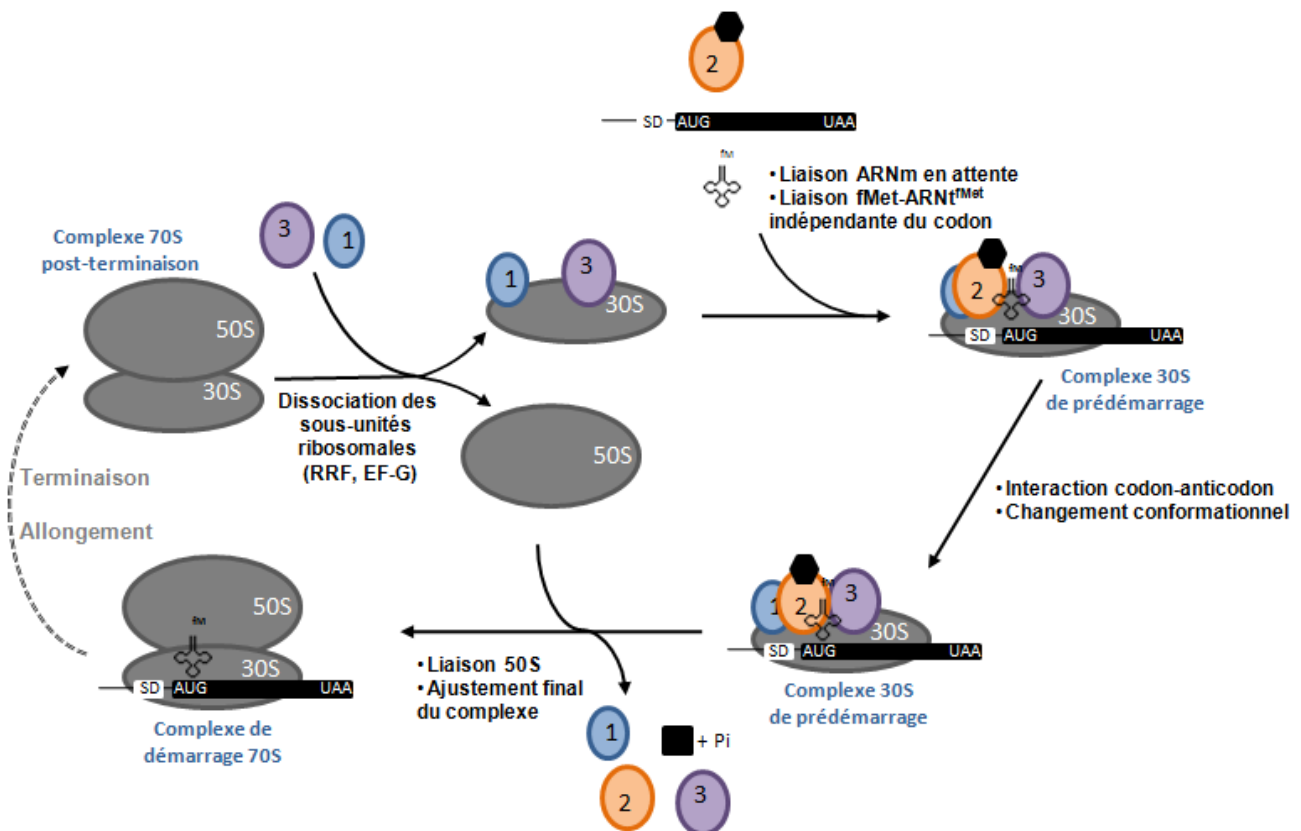


Figure 1.12 : Le démarrage de la traduction chez les Bactéries.

1 Liaison de l'ARNm

La fixation de l'ARNm chez les Bactéries est fortement dépendante d'une séquence promotrice présente dans sa région 5' non codante. Cette séquence dite Shine-Dalgarno (SD) est complémentaire d'une séquence présente en 3' de l'ARNr 16S. Chez *E. coli*, la séquence consensus GGAGG est trouvée 7 à 8 nucléotides en amont du codon de démarrage ([Shine and Dalgarno, 1974](#)). La conséquence immédiate de cette interaction est le prépositionnement du codon de démarrage à proximité du site P, lieu de l'appariement avec l'anticodon de l'ARNt initiateur. La liaison de la séquence SD se produit au niveau du cou de la sous-unité 30S, entre la tête et le dos de la plateforme. Une interaction additionnelle entre la protéine S1 et une séquence riche en pyrimidine en 5' de la séquence SD contribue à l'efficacité de la reconnaissance ([Boni et al., 1991](#)). Le facteur IF3 stimule alors l'ajustement du codon de démarrage au site P afin de favoriser l'appariement codon-anticodon ([La Teana et al., 1995](#)).

2 Liaison de l'ARNt initiateur méthionylé formylé

L'ARNt_f^{Met} possède un élément déterminant nucléotidique caractéristique. Celui-ci consiste en l'absence d'un appariement canonique entre les bases 1 et 72 alors que la plupart des ARNt élongateurs possède une paire G:C en cette position. Cette caractéristique participe à la sélection de l'ARNt initiateur par l'appareil de démarrage. Celle-ci n'affecte pas son aminoacylation par la méthionine ARNt synthétase puisque cette enzyme reconnaît essentiellement la séquence anticodon. Par contre, l'absence d'appariement C1:A72 est nécessaire pour la reconnaissance de l'ARNt initiateur par la méthionyl-ARNt_f^{Met} transformylase. L'ouverture de la paire 1:72 permet à la méthionine portée par l'ARNt initiateur de pénétrer suffisamment profondément dans le site actif de la protéine ([Guillon et al., 1993](#); [Schmitt et al., 1998](#)). Ainsi, seul l'ARNt initiateur est sélectionné par l'enzyme.

Finalement, il faut noter que la formylation de l'ARNt initiateur est absolument nécessaire au démarrage de la traduction ([Guillon et al., 1992a](#)). En effet, le groupement formyl assure sa sélection par le facteur de démarrage IF2 alors que les facteurs d'allongement ne peuvent le lier ([Guillon et al., 1996](#); [Seong and RajBhandary, 1987b](#)).

3 Le facteur IF2

Le facteur IF2 est impliqué dans la liaison de l'ARNt initiateur au ribosome. Il est doté d'un domaine de liaison du GTP et requiert la participation du centre d'activation des GTPase de la grande sous-unité pour pouvoir hydrolyser son cofacteur. Chez de nombreuses Bactéries, cette protéine existe sous plusieurs isoformes correspondant à des délétions de leurs domaines N-terminaux. Ces isoformes sont codées par un même ARNm possédant plusieurs codons de démarrage en phase. L'équilibre entre les différentes formes semble requis pour assurer une croissance optimale ([Howe and Hershey, 1983](#); [Sacerdot et al., 1992](#)). Les raisons de l'existence de cet équilibre restent encore à élucider. Ces troncatures n'affectent pas la partie C-terminale hautement conservée principalement impliquée dans l'activité du facteur. Celle-ci contient notamment le domaine G et le domaine de fixation de l'ARNt ([Spurio et al., 2000](#)). L'importance du groupement formyl est soutenue par le critère nécessaire et suffisant de sa présence pour qu'un ARNt méthionylé soit pris en charge par ce facteur. En effet, IF2 est capable de se lier à des ARNt aminoacylés dont l'acide aminé a été formylé mais ne fixe aucun ARNt ne portant pas la modification ([Sundari et al., 1976](#); [Wu and RajBhandary, 1997](#)). En outre, le nucléotide fixé au domaine G d'IF2 n'influence pas la liaison ([Pon et al., 1985](#)).

4 Liaison d'IF2 et du fMet-ARNt_f^{Met} au ribosome

La mise en évidence de l'interaction entre le facteur IF2 et l'ARNt initiateur *in vitro* a tout d'abord laissé penser à un mécanisme dans lequel IF2 fixerait l'ARNt initiateur méthionylé formylé avant de l'amener au ribosome. Cependant, la faible affinité observée en solution pour le complexe ($K_d = 1,8\mu\text{M}$) et son instabilité en présence d'ions magnésium semblent plutôt indiquer une autre séquence d'événements ([Lockwood et al., 1971](#); [Majumdar et al., 1976](#)). La spécificité pour l'ARNt initiateur au niveau du site P de la petite sous-unité ribosomale fournit déjà une base solide pour le recrutement de l'ARNt initiateur méthionine formylé. De plus, le facteur IF2 possède une affinité élevée et spécifique envers les sous-unités ribosomales. Dans un premier temps, l'ARNt initiateur est lié au site P par une voie stimulée par IF2 mais indépendante de la présence de l'appariement codon-anticodon ([Tomsic et al., 2000](#)). Le facteur IF1 stimulerait la liaison du facteur IF2 à la petite sous-unité par une interaction entre ces facteurs ([Boileau et al., 1983](#)).

5 Liaison d'IF1 au ribosome

De nombreuses études ont cherché à identifier le site de liaison d'IF1 au ribosome afin d'expliquer les effets stimulateurs observés. Les résultats biochimiques obtenus ont pu être confirmés par une structure du complexe 30S lié à IF1 obtenue par cristallographie et diffraction aux rayons X ([Carter et al., 2001](#); [Dahlquist and Puglisi, 2000](#)). Le facteur se lie au niveau du site A de la petite sous-unité. Il provoque le mouvement des résidus de l'hélice 44 et du résidu G530 dont nous avons vu l'importance. Un tel positionnement empêche ainsi toute perturbation du démarrage par fixation d'un ARNt élongateur au site A. De plus, le facteur modifie la conformation de la petite sous-unité ribosomale.

6 L'ajustement du complexe formé

Le facteur de démarrage IF3 participe également à la stabilisation de l'ARNt initiateur au site P ([Gualerzi et al., 1977](#); [O'Connor et al., 2001](#); [Weiel and Hershey, 1982](#)). Le complexe formé sur la petite sous-unité ribosomale contient alors les trois facteurs de démarrage, l'ARNm et le fMet-ARNt_f^{Met}. Il constitue le complexe de prédémarrage au sein duquel le codon de l'ARNm n'est pas apparié à l'anticodon de l'ARNt initiateur. Un changement de conformation de ce complexe va permettre de le convertir en un complexe de prédémarrage 30S stable. Celui-ci contient un fMet-ARNt_f^{Met} dont l'anticodon est positionné face à un triplet complémentaire de l'ARNm. Le facteur IF3 stimule ce changement de conformation afin de permettre un contrôle de l'appariement codon-anticodon ([Gualerzi et al., 1977](#); [Pon and Gualerzi, 1984](#)). Dans ce contexte, IF3 permet la dissociation de complexes comportant un mauvais ARNt au site P ([Gualerzi et al., 1979](#); [Hartz et al., 1990](#)) ou de complexes formés sur des codons autres que les codons canoniques du démarrage de la traduction ([Sussman et al., 1996](#)). De plus, il apparaît que le facteur pourrait agir en inspectant l'appariement codon-anticodon puisqu'il permet un démarrage sur un codon autre que AUG pourvu que l'ARNt initiateur porte un codon complémentaire ([Meinzel et al., 1999](#)).

Des études récentes visant à déterminer l'influence des différents acteurs sur les différents paramètres cinétiques caractéristiques du démarrage ont permis de préciser les étapes ultimes du processus ([Grigoriadou et al., 2007a](#); [Grigoriadou et al., 2007b](#)). Le modèle du démarrage bactérien précédent postulait qu'une fois l'appariement réalisé, les facteurs IF1 et IF3 sont éjectés du ribosome. Ces études récentes indiquent qu'IF3 pourrait rester lié à un complexe 70S intermédiaire qui ne contiendrait pas encore tous les ponts entre les sous-unités ribosomales. La conversion de ce complexe en un complexe 70S plus stable requerrait le départ du facteur. Or, la reconstitution de systèmes de traduction bactérien dénués du facteur IF3 a montré que le démarrage est plus tolérant vis-à-vis d'un codon de démarrage AUU à la place du codon canonique AUG. Cette propriété est d'ailleurs utilisée par *E. coli* pour l'autorégulation du gène codant IF3 ([Gold et al., 1984](#)). L'effet observé est attribué à la levée de

la répression de la conversion du complexe 70S en un complexe stable ([Grigoriadou et al., 2007b](#)). Ainsi, l'action d'IF3 résulte alors d'un effet déstabilisant du facteur sur le complexe 70S avant sa conversion. Une fois le complexe 70S stable formé, le fMet-ARNt^{Met} subit un ajustement ultime de son positionnement au site P. Cette étape requiert l'activité GTPasique du facteur de démarrage IF2 ([Grigoriadou et al., 2007a](#)). Le facteur pourrait ajuster l'orientation de l'ARNt initiateur au regard de la liaison de la grande sous-unité ribosomale et notamment au PTC. Finalement, le facteur IF2 est éjecté du ribosome.

Le ribosome 70S obtenu contient ainsi un fMet-ARNt^{Met} dont l'anticodon est apparié à un codon de démarrage de l'ARNm et dont le site A est prêt à accueillir un ARNt élongateur pour le premier cycle d'allongement.

II Le démarrage chez les Eucaryotes

Le jeu de facteur requis par le système de démarrage de la traduction chez les Eucaryotes dépasse de loin celui utilisé par les ribosomes bactériens et indique d'ores et déjà un processus plus complexe. Comme dans le cas bactérien, le démarrage de la traduction se déroule sur la petite sous-unité ribosomale. Ainsi, la première étape de dissociation des sous-unités est commune aux deux systèmes. La dissociation aurait également lieu par la fixation de protéines au niveau de l'interface entre les sous-unités. Parmi les facteurs de démarrage qui seraient impliqués dans la régénération de sous-unités 40S libres, le facteur de démarrage eIF6, unique au système eucaryotique, a longtemps été considéré comme un agent anti-associatif. Lié à la sous-unité 60S, ce facteur l'empêche de s'associer à la petite sous-unité ([Ceci et al., 2003](#); [Si et al., 1997](#)). Néanmoins, son rôle est certainement situé en amont de la traduction à savoir la biogénèse des ribosomes et cette protéine aurait été considérée à tort comme un facteur de démarrage ([Basu et al., 2001](#); [Sanvito et al., 1999](#); [Si and Maitra, 1999](#)).

Une extrapolation de l'implication d'IF3 dans le recyclage des ribosomes bactériens laisserait penser que le facteur eIF1, son homologue structural eucaryotique, pourrait également intervenir. Le facteur eIF3, spécifique aux Eucaryotes, en présence d'eIF1A, l'homologue eucaryotique du facteur bactérien IF1, jouerait également un rôle dans ce mécanisme. La compréhension de cette étape est limitée par le fait qu'aucun homologue eucaryotique du facteur de recyclage bactérien RRF n'a pu être identifié et les étapes ultimes de la terminaison de la traduction chez les Eucaryotes restent obscures ([Kapp and Lorsch, 2004b](#)). L'identification des acteurs participant au recyclage des ribosomes eucaryotiques devrait permettre d'éclaircir les mécanismes permettant de générer des sous-unités libres pour un nouveau cycle de traduction. Le processus de démarrage chez les Eucaryotes est schématisé sur la **figure 1.13**.

1 Liaison de l'ARNt initiateur : formation du complexe de prédémarrage 43S

L'ARNt initiateur eucaryotique présente une paire de bases A1:U72 caractéristique des Eucaryotes, ce qui le différencie des ARNt élongateurs comportant très souvent une paire G:C à cette position. De plus, contrairement au cas bactérien, la méthionine aminoacylée sur cet ARNt n'est pas modifiée. Son recrutement au ribosome apparaît aussi bien différent. Alors qu'il existe chez les Eucaryotes un facteur de démarrage fortement homologue au facteur bactérien IF2 : eIF5B, celui-ci n'assure pas les premières étapes de recrutement de l'ARNt initiateur méthionylé. En effet, cette fonction est accomplie par un facteur hétérotrimérique eIF2 possédant également une activité GTPasique mais n'ayant pas d'équivalent bactérien. A la différence du facteur IF2, cet hétérotrimère forme un complexe ternaire de forte affinité (de l'ordre du nM) avec l'ARNt initiateur méthionylé et le nucléotide GTP en solution. Le facteur eIF2 recrute ainsi l'ARNt initiateur en reconnaissant la méthionine et la paire de bases caractéristique A1:U72 et l'amène jusqu'à la sous-unité 40S.

La liaison du complexe ternaire à la petite sous-unité est fortement stabilisée par la présence des facteurs eIF3, eIF1A et eIF1. De façon intéressante, il a pu être établi chez la levure l'existence d'un complexe multifacteur composé du complexe ternaire associé aux facteurs de démarrage eIF1, eIF3 et eIF5. Ce complexe multifacteur pourrait ainsi se former en dehors du ribosome et l'ensemble serait recruté jusqu'à la petite sous-unité ([Asano et al., 2001a](#)). Le facteur eIF5 de ce complexe joue le rôle de stimulateur de l'activité GTPasique d'eIF2.

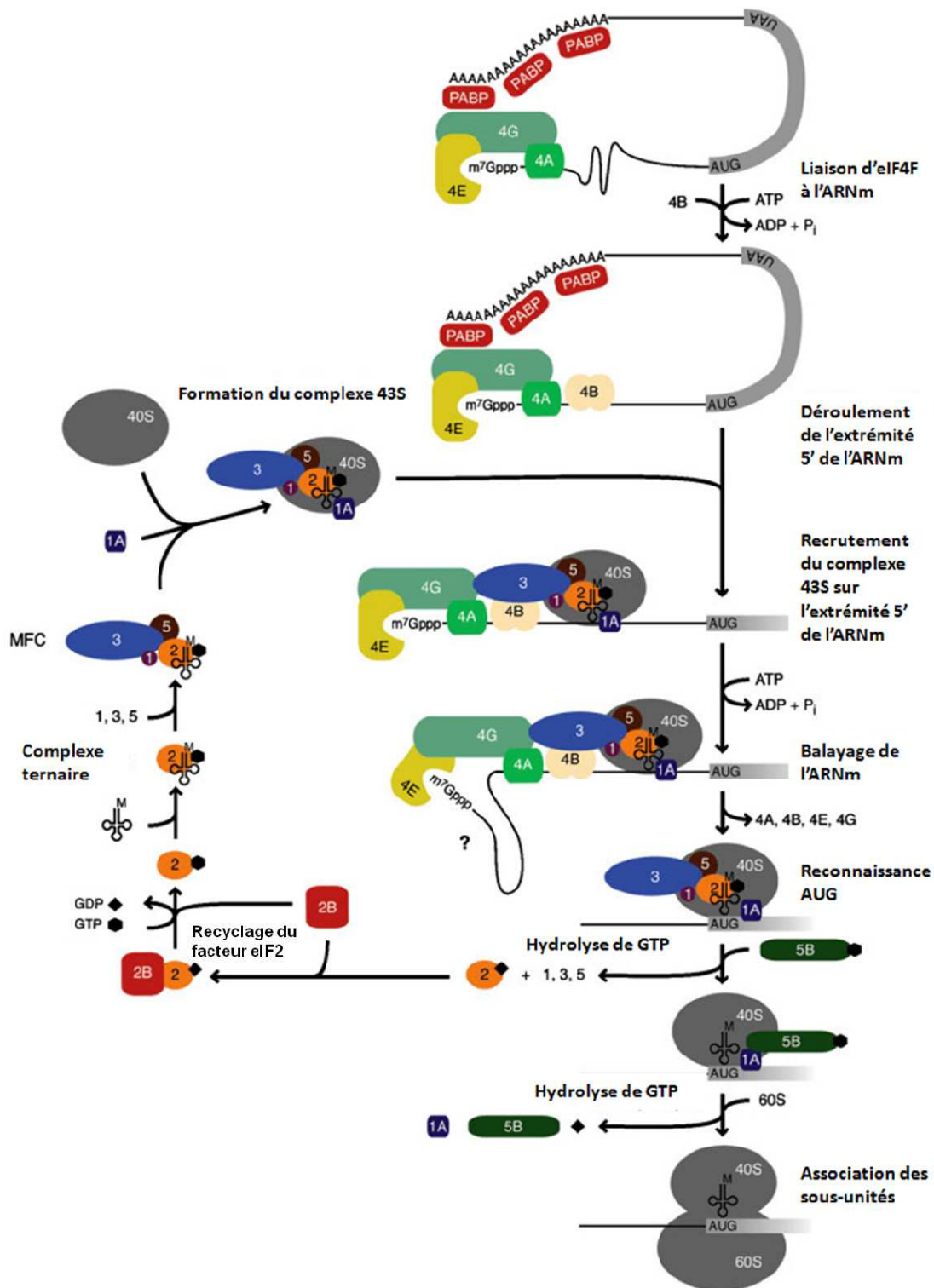


Figure 1.13 : Le démarrage de la traduction eucaryotique d'après ([Preiss and M., 2003](#)).

2 Liaison de l'ARNm : le complexe eIF4F

Le recrutement de l'ARNm chez les Eucaryotes repose sur la coopération de plusieurs protéines. La sous-unité eIF4G constitue la pièce centrale du complexe hétérotrimérique eIF4F. Les deux autres sous-unités du complexe ainsi que d'autres partenaires impliqués dans le démarrage y sont accrochées. eIF4G lie en effet la sous-unité eIF4E du complexe eIF4F mais aussi la protéine de liaison de la queue poly(A) (poly(A) Binding Protein : PABP), caractéristique des ARNm eucaryotiques. La sous-unité eIF4E est quant à elle capable de lier spécifiquement l'autre déterminant caractéristique des ARNm des Eucaryotes, à savoir la coiffe présente à l'extrémité 5'. Ainsi, ce domaine d'eIF4G organise un réseau d'interaction liant l'ARNm en 5', d'une part, et en 3', d'autre part et donc circularise l'ARNm (**Figure 1.13**). Sur le domaine central d'eIF4G se trouvent les sites de liaison d'eIF3, de la dernière sous-unité du complexe, eIF4A, et de partenaires ARN. Le facteur eIF4A possède une activité ATPasique couplée à une activité hélicase lui permettant de dérouler les éventuelles structures secondaires présentes sur l'ARNm en aval de la coiffe.

La liaison d'eIF3 à eIF4G a facilité l'assemblage du complexe de prédémarrage 43S sur l'extrémité 5' de l'ARNm.

3 La sélection du codon de démarrage : le balayage

Dans le système de démarrage des Eucaryotes, il n'existe pas d'équivalent du processus de prépositionnement du codon initiateur par une séquence de l'ARNm complémentaire de l'ARNr 16S. Le modèle de « balayage » que proposa Marilyn Kozak postule que le complexe de prédémarrage se déplace linéairement depuis l'extrémité 5' de l'ARNm jusqu'à rencontrer un codon de démarrage. Du point de vue du mécanisme, il a pu être démontré que les facteurs eIF1 et eIF1A sont nécessaires afin de permettre le mouvement de balayage : un complexe 43S formé sans ces facteurs reste en effet bloqué au niveau de la coiffe en 5' ([Pestova et al., 1998](#)). Le mécanisme de sélection du codon de démarrage sera discuté plus loin. Lorsque l'ARNt initiateur méthionylé est correctement lié, le facteur eIF2 lié au GDP est éjecté du ribosome. Le facteur étant plus affiné pour le GDP que pour le GTP, l'intervention du facteur d'échange du nucléotide guanine (Guanine Exchange Factor : GEF) : eIF2B est nécessaire. Ainsi, après chaque cycle de démarrage, le facteur est recyclé sous sa forme active liée au GTP afin de pouvoir fixer une nouvelle molécule d'ARNt.

4 L'association des sous-unités ribosomales : eIF5B

Cette étape est longtemps restée incomprise et le facteur eIF2 retenait toute l'attention pour son accomplissement, jusqu'à l'identification d'un homologue du facteur bactérien IF2, eIF5B. Une catalyse de l'étape d'association des sous-unités ribosomales par le facteur eIF5B a pu être démontrée. L'hydrolyse du cofacteur est requise pour l'éjection du facteur ([Pestova et al., 2000](#); [Shin et al., 2002](#)). A l'instar d'IF2, la liaison du facteur eIF5B sur le complexe 48S est stimulée par une interaction avec le facteur eIF1A, homologue du facteur bactérien IF1 ([Marintchev et al., 2003](#)). De façon concomitante au processus d'association des sous-unités, les facteurs eIF5B, eIF3, eIF1 et eIF1A sont éjectés du ribosome. En ce qui concerne le complexe eIF4F, bien peu d'éléments permettent d'envisager le placement du facteur après son recrutement et son devenir au cours du balayage, et même du cycle d'allongement.

III Le démarrage chez les Archées

La découverte de ces organismes dans les années 1970 a ouvert des perspectives à de multiples niveaux. Tout d'abord, d'un point de vue phylogénétique, l'absence de noyaux au sein des Archées a tout d'abord incité à les placer parmi les Bactéries. Cependant, le séquençage de génomes archéens a très vite fait apparaître des caractéristiques de ces organismes semblables à celles rencontrées dans le monde eucaryotique. La plupart des protéines archéennes trouve des homologues dans l'un ou l'autre domaine du vivant ([Woese et al., 1978](#)). Une troisième branche issue de l'ancêtre commun a donc été ajoutée à l'arbre phylogénétique afin de satisfaire ces observations ([Woese and Fox, 1977](#); [Woese et al., 1990](#)).

Dans ce contexte, l'étude de ces organismes est intéressante à deux niveaux. D'une part, les niches occupées par ces organismes sont très spécifiques et correspondent souvent à des conditions de vie extrêmes. Ces conditions peuvent être une limitation pour les études en biologie cellulaire mais représentent un avantage considérable au niveau de la biologie moléculaire. En effet, les différentes molécules des Archées doivent pouvoir résister aux conditions de vie de ces organismes. Ces propriétés facilitent bien souvent leur purification et en font des outils simples à manipuler.

D'autre part, en tant que nouvelle branche phylogénétique, la compréhension des Archées peut nous donner de nouveaux indices sur l'ancêtre commun mais aussi sur les différenciations des trois domaines de la vie. De façon remarquable lorsque l'on s'intéresse au processus de la traduction, on identifie des caractéristiques provenant à la fois des systèmes bactériens et eucaryotiques. Notamment, les ARNm archéens sont très proches des ARNm bactériens alors que le jeu de facteurs de démarrage ressemble nettement à celui mis en place par les Eucaryotes (**Tableau 1.1, p.23**).

1 Liaison de l'ARNm : un démarrage bactérien

Comme chez les Bactéries, l'absence de noyau indique que la traduction serait couplée à la transcription. De même, les ARNm archéens peuvent contenir plusieurs cistrons et la présence d'une séquence SD en amont du codon de démarrage indique un prépositionnement de l'ARNm sur la petite sous-unité ribosomale par un appariement avec une séquence complémentaire de l'ARNr 16S. Il est cependant à noter que certaines Archées possèdent de nombreux ARNm sans séquence SD dont le mode de démarrage correspondrait à un héritage de celui utilisé par l'ancêtre universel commun. Les caractéristiques bactériennes des ARNm archéens sont en bon accord avec l'absence du complexe eIF4F et des facteurs eIF4A, eIF4B et eIF4H chez ces organismes.

2 Formation d'un complexe 80S : un démarrage eucaryotique

L'ARNt initiateur archéen est très semblable à celui des Eucaryotes. En effet, il contient la paire de bases conservée chez les Eucaryotes A1:U72 et la méthionine ne présente pas de modification. De plus, le recrutement de l'ARNt à la petite sous-unité ribosomale fait intervenir le facteur aIF2, homologue au facteur eIF2. En outre, les facteurs eIF1 et eIF1A que nous avons vus impliqués dans le processus de sélection du codon de démarrage chez les Eucaryotes trouvent des homologues dans ce système : aIF1 et aIF1A respectivement (**Tableau 1.1, p.22**).

Néanmoins, l'absence d'homologue fonctionnel et structural pour eIF3 soulève quelques questions. La présence du facteur aIF2 nous inciterait à chercher des homologues de ses partenaires, à savoir le GEF hétéropentamérique eIF2B et la GAP (Guanine Activating Protein) eIF5. Or, seules trois sous-unités du facteur aIF2B ont été identifiées chez les Archées. Parmi celles-ci, on ne retrouve pas d'homologue de la sous-unité eucaryotique essentielle à la catalyse du recyclage d'aIF2. Ainsi, l'éventualité d'un recyclage spontané chez ces organismes

reste la plus probable. De plus, il n'existe pas d'homologue du facteur activateur eIF5. On peut envisager qu'un facteur ne possédant pas d'homologie de séquence avec eIF5 remplisse ce rôle ou bien qu'eIF2 possède intrinsèquement une activité GTPasique suffisante. Enfin, l'absence d'un homologue du facteur eIF3 surprend au vu du rôle central d'adaptateur que ce complexe réalise. Il faut cependant remarquer que ce facteur organise notamment un réseau d'interaction avec le complexe eIF4F afin de recruter l'ARNm sur la sous-unité 30S. La voie de sélection de l'ARNm plus proche de celle des Bactéries pourrait donner les premiers éléments de réponse quant à l'absence de ce facteur.

La description du démarrage dans les trois domaines du vivant révèle de façon très nette des caractères hérités de l'ancêtre universel commun. L'état des connaissances permet de dépeindre la plupart des mécanismes élémentaires. La suite de cette introduction se focalisera sur les relations structure-fonction des éléments du système de démarrage des Eucaryotes et en particulier chez la levure *Saccharomyces cerevisiae*.

PARTIE C RELATIONS STRUCTURE/FONCTION DES FACTEURS DU DEMARRAGE EUCARYOTIQUE

I Le complexe de prédémarrage 43S

1 Les composants du complexe 43S

a Le complexe ternaire $eIF2:GTP:Met-ARNt_i^{Met}$

Le recrutement de l'ARNt initiateur au ribosome requiert la formation préalable d'un complexe ternaire comportant le facteur de démarrage eIF2, le GTP et le Met-ARNt_i^{Met}. Le facteur de démarrage eIF2 est un complexe hétérotrimérique composé des sous-unités α (36kDa), β (38kDa) et γ (58kDa) chez la levure. Le facteur eucaryotique est très homologue au facteur archéen sur lequel nous possédons de nombreux éléments biochimiques et structuraux. Une différence notable réside néanmoins dans l'absence chez les Archées d'une extension N-terminale de la sous-unité β (**Figure 1.14**). Le domaine spécifique aux Eucaryotes comporte notamment trois groupes de résidus lysine ou boîtes K impliquées dans la liaison des partenaires de l'activité GTPase eIF5 et du recyclage eIF2B (sous-unité ϵ) ([Asano et al., 1999](#); [Das and Maitra, 2000](#)). L'absence de ces partenaires dans le système archéen est sans doute reliée à l'absence de leur domaine de liaison chez aIF2 β ([Kyrpides and Woese, 1998a](#)).

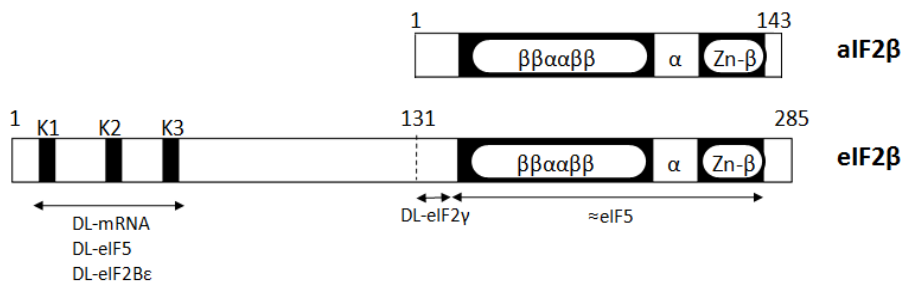


Figure 1.14 : Domaines importants des facteurs eIF2 β (numérotation de *S. cerevisiae*) et aIF2 β (numérotation *Pyrococcus abyssi*). Les doubles flèches indiquent les domaines de liaison (DL) de l'ARNm ou des facteurs de démarrage et la partie homologue (\approx) à eIF5. Les motifs polylysines (K), le domaine de liaison du zinc (Zn- β) et les repliements de la protéine sont indiqués.

Structuralement, la sous-unité γ du facteur archéen constitue le cœur du complexe hétérotrimérique : les deux autres sous-unités sont associées avec cette sous-unité mais n'interagissent pas entre elles. De plus, la sous-unité γ contient les sites de liaison du cofacteur et du Met-ARNt_i^{Met} ([Schmitt et al., 2002](#)). La résolution de structures par cristallographie aux rayons X de la sous-unité γ de l'Archée *Pyrococcus abyssi* révèle une homologie structurale forte avec les facteurs d'allongements bactérien et eucaryotique, EF1A et eEF1A respectivement (**Figure 1.15**). Ces GTPases assurent la liaison et le recrutement au ribosome des ARNt aminoacylés au cours de l'allongement. aIF2 γ s'organise en trois domaines, le premier domaine contient le site de liaison du cofacteur (domaine G) ; les deux autres domaines ont des structures en tonneau β . Par analogie avec les facteurs d'élongation, un canal de liaison de l'extrémité 3' du Met-ARNt_i^{Met} a pu être déduit. Cette poche de liaison conservée présente une surface chargée positivement pouvant accommoder le squelette polyphosphaté et les bases de la tige acceptrice. Dans la famille des protéines G, l'activité GTPasique est couplée à un mouvement de 2 régions caractéristiques appelées « switch » 1 et 2. En état « switch on », des liaisons hydrogènes entre des résidus de chacune de ces régions et le phosphate γ du GTP laissent la poche de liaison de la tige acceptrice de l'ARNt vacante. Lorsque le phosphate est perdu, la relaxation

des switches obstrue cette poche de liaison, ce qui diminue fortement l'affinité du facteur pour le Met-ARNt_i^{Met} (Yatime et al., 2006). L'hydrolyse du GTP et l'éjection du facteur seront discutées par la suite.

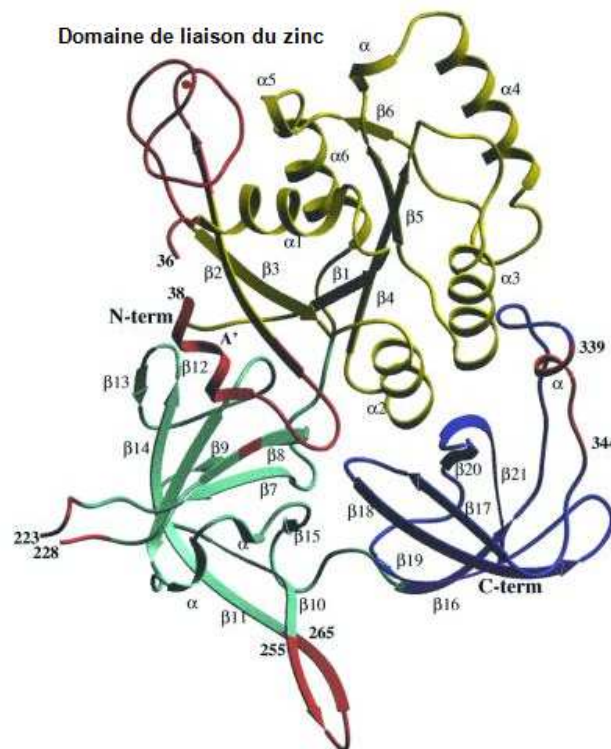


Figure 1.15 : Structure tridimensionnelle d' α IF2 γ de *P. abyssi* d'après (Schmitt et al., 2002). Le domaine I est coloré en jaune, le domaine II en bleu turquoise et le domaine III en bleu foncé. Le domaine de liaison du zinc est représenté en rouge. Les autres régions caractéristiques du facteur apparaissent en rouge et notamment la région switch 1 comprenant l'hélice A' (résidus 32-50).

La structure de la sous-unité α d'Archée, et celles de sous-unités α eucaryotiques révèlent une structure en deux domaines mobiles l'un par rapport à l'autre (Ito et al., 2004; Yatime et al., 2006, 2007; Yatime et al., 2005) (Figure 1.16). Le domaine N-terminal se compose d'un sous-domaine présentant un repliement en tonneau β de type OB-fold (Oligonucléotide/Oligosaccharide Binding) fréquent dans les protéines liant des acides nucléiques. Il interagit avec un second sous-domaine hélicial. Le domaine OB-fold permet effectivement une interaction avec l'ARN mais celle-ci semble non spécifique (Yatime et al., 2004). Le domaine C-terminal adopte une structure en α/β . La structure du dimère $\alpha\gamma$ puis du facteur $\alpha_3\beta\gamma$ (où α_3 correspond à la partie C-terminale de la sous-unité) montrent que la sous-unité α se lie au deuxième domaine de γ proche de la poche de liaison de l'ARNt, mais qu'une interaction directe avec l'ARNt est difficilement envisageable (Yatime et al., 2006, 2007). Les études biochimiques menées sur le facteur archéen indiquent que la sous-unité α est impliquée dans la liaison de la sous-unité γ au Met-ARNt_i^{Met} et que son domaine C-terminal (α_3) suffit pour réaliser cette stimulation (Yatime et al., 2004, 2005). L'effet observé serait donc indirect et pourrait être lié à la stabilisation des régions switch par la sous-unité α (Yatime et al., 2006). De plus, chez les Archées, la sous-unité β ne modifie pas cette liaison dépendante de la présence de GTP et de la paire de base discriminatrice A1:U72 de l'ARNt initiateur (Kapp and Lorsch, 2004a; Yatime et al., 2004). Les données disponibles sur eIF2 de levure indiquent plutôt que le rôle stimulateur pour la liaison du facteur à l'ARNt initiateur serait accompli par la sous-unité β alors que la sous-unité α n'aurait pas d'influence (Nika et al., 2001).

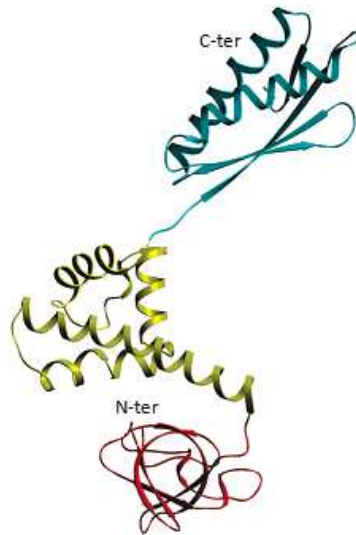


Figure 1.16 : Structure tridimensionnelle d'eIF2 α de *P. abyssi* d'après (Yatime et al., 2005). Les sous-domaines N-terminaux sont représentés en rouge (OB-fold) et jaune (sous-domaine hélical). Le domaine C-terminal apparaît en cyan.

Le domaine C-terminal d'eIF2 β est conservé chez les Eucaryotes et les Archées (Figure 1.14). La structure tridimensionnelle de la sous-unité β archéenne par résonance magnétique nucléaire (RMN) indique une architecture en deux domaines de la partie conservée. Un domaine de liaison du zinc nécessaire à l'activité (ou ZBD : Zinc Binding Domain) prolonge un domaine central structuré en motif α/β . L'extension N-terminale des Eucaryotes apparaît non structurée (Cho and Hoffman, 2002; Gutierrez et al., 2002). La résolution des structures du dimère $\beta\gamma$ et du trimère $\alpha_3\beta\gamma$ archéens permettent de visualiser l'hélice N-terminale liée au domaine G de la sous-unité γ (Sokabe et al., 2006; Yatime et al., 2007) (Figure 1.17). Cependant, ces deux structures font apparaître des conformations différentes de la sous-unité β . Un mouvement rigide des domaines ZBD et α/β par rapport à l'hélice N-terminale est observé (Figure 1.17).

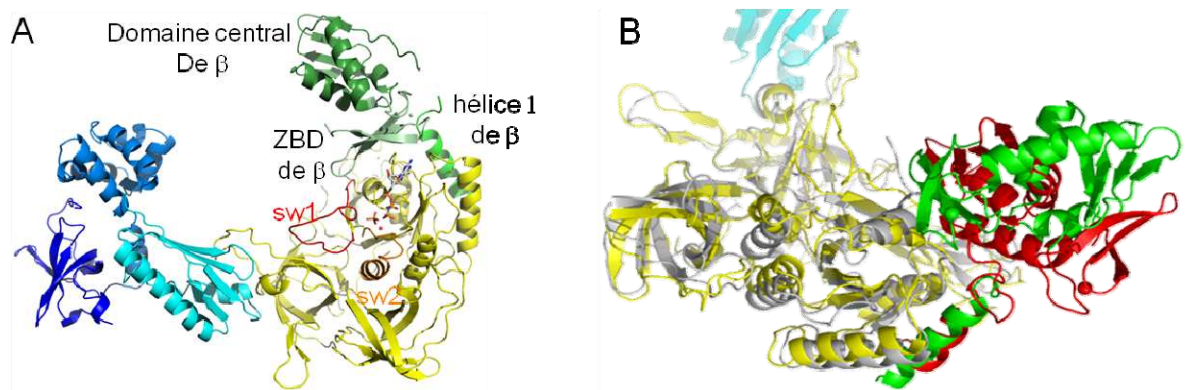


Figure 1.17 : Structure de l'hétérotrimère eIF2 de *Sulfolobus solfataricus* et comparaison au dimère $\beta\gamma$ de *P. abyssi* d'après (Yatime et al., 2007). (A) Représentation du trimère de *S. solfataricus* obtenue en superposant la sous-unité α_3 (domaine C-terminal de α) du trimère $\alpha_3\beta\gamma$ cristallisé avec la structure de la sous-unité α seule (Yatime et al., 2006). Les sous-unités α et γ sont représentées en bleu et en jaune, respectivement. Les régions switch 1 et 2 de la sous-unité γ sont indiquées en rouge (SW1) et orange (SW2), respectivement. La sous-unité β apparaît en vert et le domaine central et le domaine de liaison du zinc (ZBD) sont indiqués ainsi que l'hélice N-terminale assurant la liaison à eIF2 γ . (B) Superposition des structures du trimère de *S. solfataricus* et du dimère $\beta\gamma$ de *P. abyssi* (Sokabe et al., 2006). L'alignement est basé sur les sous-unités γ . Le trimère apparaît dans le même code couleur que (A). Les sous-unités β et γ du dimère sont représentées respectivement en rouge et en gris.

b Le facteur eIF5

Cette protéine partenaire du facteur eIF2 l'assiste en stimulant son activité GTPasique lorsque le facteur est engagé dans un complexe 48S ([Das et al., 2001](#); [Das and Maitra, 2000](#)). La partie N-terminale du facteur eIF5 contient des domaines structurellement homologues aux domaines N-terminal et ZBD d'eIF2 β ([Cho and Hoffman, 2002](#); [Conte et al., 2006](#)) (**Figure 1.18**). La partie C-terminale contient des boîtes AA (Acides Aromatiques). D'une part, les boîtes AA de la région C-terminale lient la région N-terminale d'eIF2 β contenant les boîtes K ([Asano et al., 1999](#); [Das and Maitra, 2000](#)). D'autre part, des travaux récents indiquent une interaction directe du domaine N-terminal d'eIF5 avec eIF2 γ chez la levure ([Alone and Dever, 2006](#)).

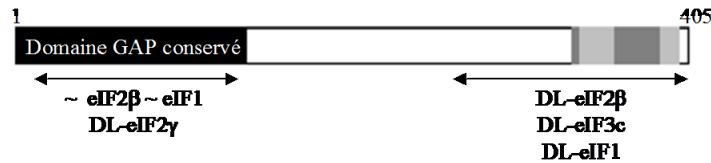


Figure 1.18 : Domaines importants du facteur eIF5 (numérotation de *S. cerevisiae*). Les doubles flèches indiquent les domaines de liaison (DL) des facteurs de démarrage et les parties homologues (~) à eIF1 et eIF2 β sont représentées. Le domaine HEAT apparaît en gris foncé et les boîtes AA en gris clair.

La structure cristallographique du domaine C-terminal présente un repliement de type HEAT ([Wei et al., 2006](#); [Yamamoto et al., 2005](#)) (**Figure 1.19**). Les domaines HEAT (Huntington; Elongation factor 3; A subunit of protein phosphatase 2A; Target of rapamycin) se composent de répétitions d'un motif constitué de deux hélices antiparallèles et deux boucles. Ces domaines sont identifiables par la présence de résidus hydrophobes et des résidus proline, arginine et aspartate conservés. Les motifs adjacents s'empilent parallèlement les uns aux autres notamment par des interactions entre aspartates et arginines conservés. Cela crée une double couche d'hélices α s'arrangeant telle une superhélice droite. La structure présente trois surfaces chargées distinctes. La première présente les résidus des boîtes AA impliquées dans la liaison de boîtes K. La deuxième surface, chargée positivement, présente de nombreuses lysines impliquées dans la liaison des facteurs eIF1 et eIF3. Enfin, une troisième zone chargée très négativement pourrait fournir une nouvelle surface d'interaction avec des protéines contenant des boîtes K ([Wei et al., 2006](#)).

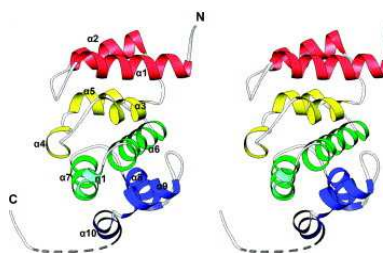


Figure 1.19 : Structure tridimensionnelle du domaine C-terminal d'eIF5 d'après ([Wei et al., 2006](#)). Dans le motif HEAT canonique, les hélices présentent un angle de 24° entre leurs axes respectifs. De plus, les hélices présentent souvent un coude de 45° induit par la présence des prolines conservées. La courbure des structures s'explique par de faibles rotations d'environ 20° entre deux répétitions consécutives selon un axe parallèle à celui des hélices. Le domaine HEAT d'eIF5 contient 10 hélices α qui forment 4 motifs de répétitions d'hélices antiparallèles. Le premier motif est canonique alors que les trois suivants sont constitués de trois hélices dont une plus petite. De façon inhabituelle, entre tous les motifs adjacents, sont observées des rotations de 20-25° parallèlement à l'axe de la superhélice.

c Le facteur eIF1A

La protéine eIF1A de 17kDa chez les Eucaryotes fait partie du cœur universellement conservé de la traduction (Kyripides and Woese, 1998b). L'homologie constatée en comparant les structures tridimensionnelles des facteurs bactérien IF1, eucaryotique eIF1A et archéen aIF1A appuie l'universalité du facteur (Battiste et al., 2000; Li and Hoffman, 2001; Sette et al., 1997) (Figure 1.20). Le facteur comprend un domaine OB-fold encadré par un domaine hélical et des extensions N et C-terminales non structurées (random coil) contenant des groupes de régions acides ou basiques importantes pour la liaison d'autres facteurs. Notamment, une extension C-terminale spécifique aux Eucaryotes supérieurs lui permet de se lier au facteur eIF5B (Marintchev et al., 2003).

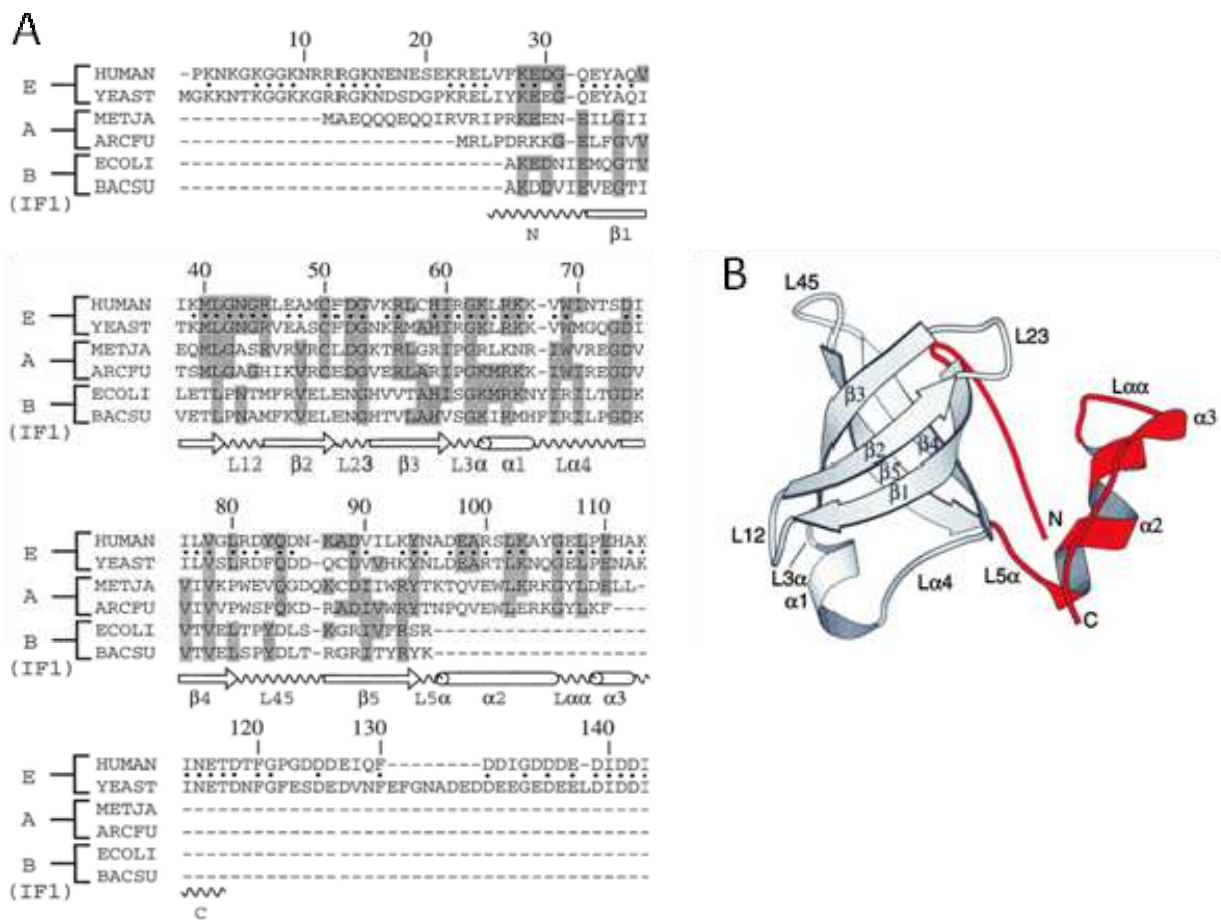


Figure 1.20 : Conservation du facteur IF1/eIF1A et structure du facteur humain d'après (Battiste et al., 2000). (A) Alignement de séquences adapté de (Kyripides and Woese, 1998b). Les acides aminés ombrés correspondent aux résidus conservés à plus de 50% dans les 12 séquences analysées par Kyripides. Les résidus conservés du facteur humain et du facteur de levure sont indiqués par des cercles. Les lettres A, E et B font références aux trois domaines du vivant. Les séquences des organismes *Homo sapiens* (HUMAN), *S. cerevisiae* (YEAST), *M. janaschii* (METJA), *Archaeoglobus fulgidus* (ARCFU), *E. coli* (ECOLI) et *Bacillus subtilis* (BACSU) sont présentées. Les structures secondaires du facteur eIF1A humain sont reportées sous les séquences. (B) Structure du facteur eIF1A humain. Le repliement OB-fold et le domaine hélical sont respectivement représentés en gris et rouge.

d Le facteur eIF1

Aucun orthologue du facteur bactérien IF3 n'avait été identifié par alignement de séquence malgré un rôle crucial de ce facteur au cours du démarrage. Le facteur eIF1, le plus petit de la machinerie eucaryotique (12kDa), semblait néanmoins partager une homologie fonctionnelle avec le facteur IF3 pour la sélection du codon de démarrage ([Castilho-Valavicius et al., 1990](#); [Pestova et al., 1998](#); [Yoon and Donahue, 1992](#)). La résolution de la structure tridimensionnelle du facteur eIF1 humain a apporté un élément important ([Fletcher et al., 1999](#)). En effet, la forte homologie structurale entre les structures du domaine C-terminal fonctionnel du facteur IF3 (noté IF3C) ([Garcia et al., 1995](#)) et du facteur eIF1 ([Fletcher et al., 1999](#)) incite à inclure dans le cœur universellement conservé un facteur de démarrage IF3C/e/aIF1. Le facteur forme une structure compacte organisée en α/β fortement homologue aux structures des domaines N-terminaux d'eIF2 β et d'eIF5 (**Figure 1.21**). Cela a conduit à définir un nouveau type de repliement (eIF125) indiquant une possible apparition de ces facteurs par duplication de gènes ([Conte et al., 2006](#)).

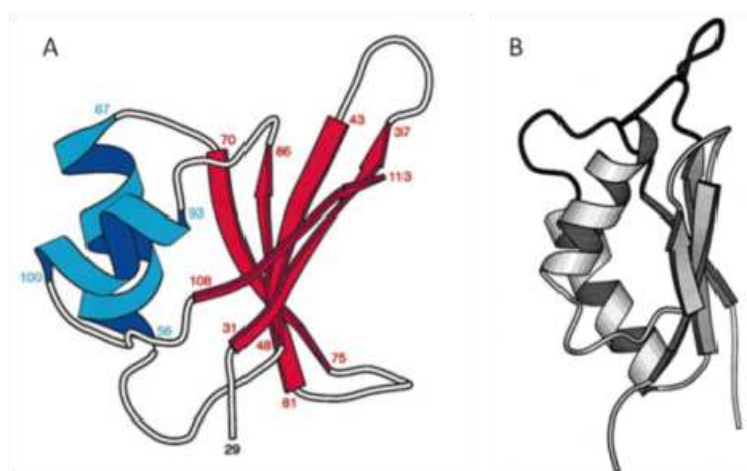


Figure 1.21 : Homologie structurale entre les facteurs eIF1 humain et IF3 bactérien. (A) Structure tridimensionnelle du facteur eIF1 humain d'après ([Fletcher et al., 1999](#)). Les structures hélicales et en feuillets β sont représentées en bleu et rouge, respectivement. (B) Structure du domaine C-terminal du facteur IF3 bactérien d'après ([Garcia et al., 1995](#)).

e Le facteur eIF3

Le facteur eIF3 est l'élément le plus complexe de la machinerie eucaryotique d'une masse molaire de plus de 650kDa chez l'humain. Ce facteur comporte 13 sous-unités chez les mammifères appelées eIF3a à m. 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). Ce complexe définit le cœur fonctionnel du facteur. Au cours du démarrage, eIF3 organise un vaste réseau d'interactions avec de nombreux partenaires parmi lesquels : eIF1, eIF1A, eIF2, eIF5, eIF4G, eIF4B et la sous-unité 40S. Une autre sous-unité non-essentielle : eIF3j stimule la liaison d'eIF3 à ses partenaires ([Fraser et al., 2004](#)). La formation de ce réseau d'interactions sera décrite dans la partie consacrée au facteur eIF3.

La coordination de ces différentes interactions en fonction du stade du démarrage reste à préciser. L'étude du système de démarrage chez la levure a néanmoins permis de mettre en évidence un complexe multifacteur comportant le complexe eIF3 sur lequel se fixent eIF1, eIF1A, eIF5 et le complexe ternaire (eIF2:GTP:Met-ARNT_i^{Met}). L'ARNt initiateur est probablement recruté sous cette forme sur la petite sous-unité ribosomale ([Asano et al., 2000](#)). La combinaison de nombreuses études biochimiques a permis de proposer un modèle d'interaction entre les différents partenaires ([Valasek et al., 2003](#)). Ces études n'ont pas identifié le site de liaison du facteur eIF1A. Cependant, l'interaction de ce facteur avec eIF2 et eIF3 observée *in vivo* indique très certainement sa présence au sein du MFC ([Olsen et al., 2003](#)).

2 Génération de sous-unités libres

Cette étape nécessaire vise à dissocier des ribosomes dits de post-terminaison à savoir liés à l'ARNm et un ARNt déacylé présent au site P. Cependant, il n'a pu être identifié de facteurs protéiques assistant ce processus. L'étape de post-terminaison crée une sous-unité 60S libre et une sous-unité 40S associée à des facteurs et notamment eIF3 ([Freiinstein and Blobel, 1975](#)). La reconstitution *in vitro* de ribosomes a confirmé l'implication du facteur eIF3 ([Kolupaeva et al., 2005](#)). Un modèle d'interaction du facteur humain sur la sous-unité 40S déduit d'expériences de cryomicroscopie électronique situe le facteur dans une région essentiellement accessible au solvant ([Siridechadilok et al., 2005](#)) (**Figure 1.22**). Cette position concorde avec les quelques études biochimiques existantes chez la levure proposant un positionnement du complexe sur la sous-unité ([Valasek et al., 2003](#)).

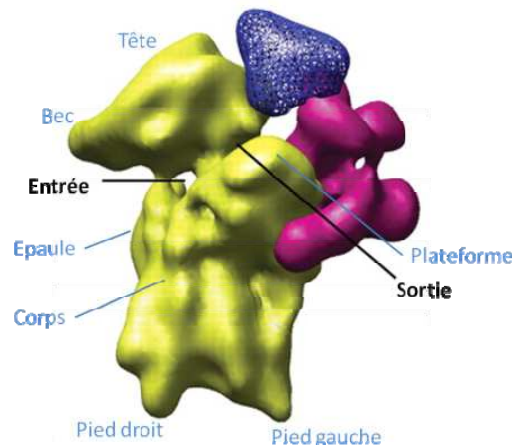


Figure 1.22 : Structure de la sous-unité 40S liée à eIF3 et eIF4G d'après ([Siridechadilok et al., 2005](#)). Ce complexe est déduit de la combinaison des structures 40S:eIF3:IRES et eIF3:eIF4G. Le ribosome est représenté en jaune, le facteur eIF3 en violet et le facteur eIF4G en violet sombre. Les différentes parties de la sous-unité 40S sont indiquées ainsi que les canaux d'entrée et de sortie de l'ARNm.

Ainsi, le facteur eIF3 n'agirait pas directement en bloquant l'interface entre les sous-unités mais plutôt en proposant des sites d'interactions pour d'autres partenaires. L'implication des facteurs eIF1 et eIF1A a été supposée assez tôt par comparaison avec le système bactérien. Une étude par coupure dirigée avec des radicaux hydroxyles fournit des informations quant au positionnement du facteur eIF1 sur la sous-unité 40S ([Fletcher et al., 1999](#); [Lomakin et al., 2003](#)). Ce modèle concorde parfaitement avec celui déduit de la position de son facteur homologue structural bactérien IF3 sur la sous-unité 30S ([Dallas and Noller, 2001](#)) (**Figure 1.23**). Le facteur se lie au niveau de l'ARNr 18S engagé dans la formation de ponts dans la région centrale d'interaction. De plus, à l'instar du facteur IF1, le facteur eIF1A pourrait également se lier au niveau du site A de la petite sous-unité ribosomale et induire les mêmes changements conformationnels ([Battiste et al., 2000](#); [Carter et al., 2001](#)) (**Figure 1.23**).

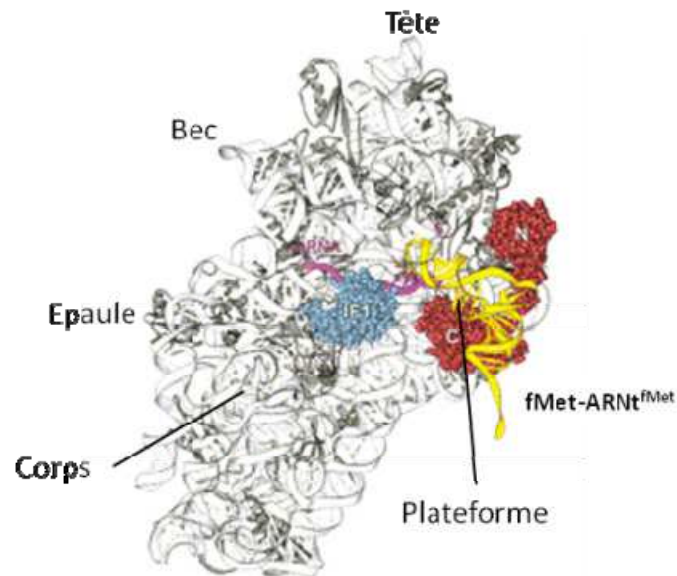


Figure 1.23 : Structure de la sous-unité 30S liée à IF1 et IF3 d'après (Dallas and Noller, 2001). Cette structure est obtenue en reportant les données biochimiques concernant la liaison d'IF3 au ribosome sur la structure cristallographique de la sous-unité 30S liée au facteur de démarrage IF1 (Carter et al., 2001). Le facteur IF1 est représenté en bleu, IF3 en rouge et ses parties N et C-terminales sont indiquées. Les positions des ligands ARNm (violet) et fMet-ARNt^{Met} (jaune) sont modélisés.

Néanmoins, l'intervention de facteurs additionnels tel un ARNm ou le complexe ternaire eIF2:GTP:Met-ARNt^{Met} semble requise (Kolupaeva et al., 2005). De plus, l'identification du complexe multifacteur (MFC) chez la levure laisse envisager que l'ensemble de ce complexe pourrait en fait être recruté à la sous-unité 40S (Asano et al., 2000; Valasek et al., 2003). L'effet combinerait alors la gêne stérique induite par les facteurs occupant l'interface de la sous-unité 40S qui lie la 60S et une modification de la conformation de la sous-unité 40S. Une caractérisation plus aboutie des dernières étapes de la traduction eucaryotique devrait permettre de mieux comprendre quels éléments sont requis *in vivo* pour recycler le ribosome.

3 Assemblage du complexe de prédémarrage 43S

De nombreuses études ont caractérisé la liaison du complexe ternaire sur la sous-unité 40S menant à la formation d'un complexe de prédémarrage 43S. Cette étape est stimulée par les facteurs eIF1A, eIF1 et eIF3 (Benne and Hershey, 1978; Chaudhuri et al., 1999; Majumdar et al., 2003; Trachsel et al., 1977). Comme nous venons de le voir, la caractérisation des facteurs présents sur la sous-unité 40S suite au recyclage du ribosome est très limitée. La chronologie des liaisons du complexe ternaire et des facteurs stimulant son recrutement ne peut alors être prédite. Le processus implique certainement des liaisons entre facteurs de démarrage, l'interaction directe entre eIF2 et la sous-unité 40S et des changements de conformation de la petite sous-unité ribosomale. Chez la levure, le complexe ternaire pourrait être recruté au sein du MFC au moment de la dissociation des sous-unités ribosomales (Asano et al., 2000; Das et al., 1997; Olsen et al., 2003).

Au sein du complexe de prédémarrage 43S, le facteur eIF3 serait lié sur la surface accessible au solvant, eIF1A occuperait le site A et le site de fixation d'eIF1 serait situé entre la plateforme et le site P, au niveau de l'interface entre les sous-unités (Figure 1.22 et 1.23). Le positionnement du facteur eIF2 reste quant à lui ambigu.

II Recrutement de l'ARNm au complexe de prédémarrage 43S et balayage de l'ARNm

Le complexe de prédémarrage 43S ainsi formé doit alors recruter l'ARN messenger et sélectionner le codon de démarrage. Dans un premier temps, le ribosome se lie à la coiffe présente en 5' de la plupart des ARNm eucaryotiques. Cette étape nécessite l'intervention coordonnée des facteurs de démarrage : eIF4F, eIF4B, eIF4H et de la protéine PABP. Leur action permettrait de dérouler les éventuelles structures secondaires de l'ARNm et de rendre sa conformation optimale pour sa liaison au complexe 43S. Ils se lient à la coiffe en 5', d'une part, et la queue poly(A) en 3', d'autre part. L'ARNm se lie autour du cou de la petite sous-unité eucaryotique, maintenu par une interaction de part en part avec les hélices 16 et 34, respectivement de l'épaule et de la tête [\(Spahn et al., 2004b\)](#). *In vitro*, il reste possible d'obtenir des complexes de prédémarrage avec seulement la petite sous-unité, eIF1, eIF1A, eIF2, l'ARNt initiateur méthionylé et l'ARNm, à condition que la région 5' non-codante ne soit pas structurée [\(Pestova and Kolupaeva, 2002\)](#).

1 Liaison de la coiffe en 5'

Le coiffage des ARNm est réalisé grâce à l'action de trois enzymes : une ARN 5'-triphosphatase, une guanylyl transferase et une N⁷G-méthyltransférase. Au cours des toutes premières étapes de la transcription, une guanosine méthylée en position 7 (m⁷G) est ajoutée à l'extrémité 5' du transcrit par un pont triphosphate 5'-5' (notée m⁷GpppN où N est le premier nucléotide de l'ARNm). La base N ainsi que des modifications additionnelles de cette structure modulent l'efficacité de la traduction des ARNm [\(Shatkin, 1976\)](#).

La reconnaissance de cette coiffe est assurée par une sous-unité du complexe eIF4F : eIF4E. L'importance de cette sous-unité se retrouve au niveau de sa séquence très conservée au sein des Eucaryotes. Les structures tridimensionnelles d'eIF4E issues de différents organismes ont révélé un repliement nouveau. La structure consiste en un feuillet β de 8 brins anti-parallèles incurvé derrière lequel 3 longues hélices sont empilées [\(Marcotrigiano et al., 1997; Matsuo et al., 1997; Tomoo et al., 2002\)](#).

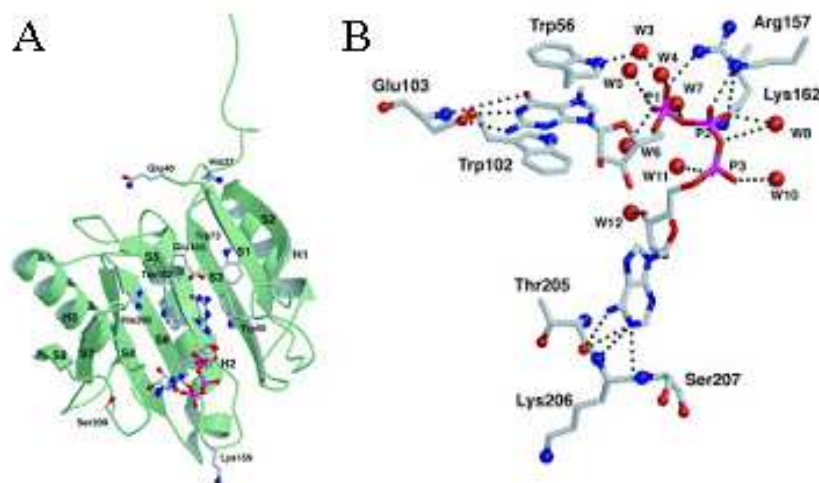


Figure 1.24 : Liaison de la coiffe par le facteur eIF4E d'après [\(Tomoo et al., 2003\)](#). (A) Structure tridimensionnelle du complexe eIF4E:m⁷GpppA. Les hélices (H) et les brins (S) sont indiqués ainsi que les résidus impliqués dans la liaison de la coiffe. (B) Vue de la poche de liaison de la coiffe. Les molécules d'eau (W) sont représentées par des boules rouges. Les liaisons hydrogène et contacts de Van der Waals possibles sont indiqués par des traits pointillés. Les phosphates de la coiffe sont indiqués P1, P2 et P3.

La coiffe est maintenue dans la partie concave du feuillet par deux interactions d'empilement π - π avec les groupements indoles des résidus hautement conservés Trp56 et 102 (la numérotation fait référence au facteur humain) (**Figure 1.24**). Ces interactions d'empilement ainsi que les phosphates de la coiffe sont stabilisés par des liaisons hydrogènes. La boucle C-terminale (Ala203-Thr211) s'organise lorsque le facteur est cocrystallisé avec un analogue m^7GpppA . L'adénine est stabilisée par des liaisons hydrogènes avec les résidus Thr205 et Thr211 et des contacts de Van der Waals avec la Ser207 (**Figure 1.24**). Cette différence dans les structures est cohérente avec l'affinité moindre d'eIF4E pour l'analogue m^7GpppG comparé à m^7GpppA ([Zuberek et al., 2003](#)).

2 Liaison de la queue poly(A) en 3'

La coiffe greffée en 5' des ARNm stimule l'ajout d'un polymère d'adénine par la poly(A) polymérase en 3' de l'ARNm. Cette extension de longueur variable peut atteindre 250 à 300 nucléotides chez les mammifères. Les ARNm peuvent être synthétisés avec une queue poly(A) relativement courte afin d'être disponibles dans le cytoplasme mais ils ne seront pas traduits efficacement. L'ajout ultérieur d'adénines stimulera leur traduction. Au contraire, la deadénylation d'un ARNm le dirigera vers une traduction moins efficace puis vers sa dégradation.

Une protéine cytoplasmique et nucléaire se lie à cette extrémité 3' : la PABP (Poly(A) Binding Protein) et agit à de multiples niveaux. La PABP contient 4 domaines RRM (RNA Recognition Motif) séparés par des domaines hautement conservés suivis d'un domaine C-terminal interagissant avec des partenaires protéiques. Sa fonction en relation avec le démarrage de la traduction réside dans sa partie N-terminale comprenant les deux premiers motifs RRM. La structure tridimensionnelle de ces deux motifs liés à un polymère de 11 adénines indique les déterminants de cette liaison ([Deo et al., 1999](#)). Chaque RRM est organisé en feuillet β derrière lequel 4 hélices α sont empilées (**Figure 1.25**). La nécessité de la présence de deux RRM en tandem est soutenue par des interactions observées entre ces motifs. Le polymère d'adénines adopte une structure étendue stabilisée par un réseau de liaisons hydrogènes, de contacts de Van der Waals et des interactions d'empilement π - π assurées par des résidus hautement conservés des RRM mais également du domaine situé entre ces motifs. La création de ce réseau requiert environ 27 résidus entre deux PABP adjacents. Ainsi, les ARNm eucaryotiques peuvent contenir entre 2 et 10 partenaires sur leurs extrémités 3'.

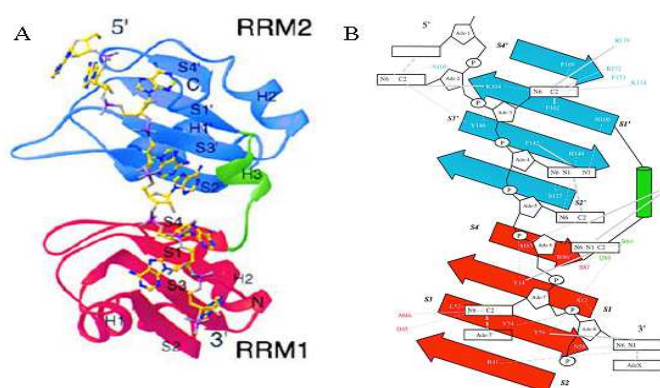


Figure 1.25 : Liaison d'un polymère d'adénine par les RRM1 et 2 de la PABP d'après ([Deo et al., 1999](#)). (A) Structure tridimensionnelle du complexe RRM1/2:ARN de la PABP humaine. Les hélices (H) et feuillets (S) des RRM1 et 2 sont indiqués (annotation ' pour le RRM2). Les N et C-terminaux de la protéine ainsi que les extrémités 5' et 3' de l'ARN sont repérées. (B) Représentation schématique des interactions ARN-protéine. Le code couleur pour les motifs RRM sont les mêmes que ceux de (A). Les traits épais indiquent les contacts Van der Waals ; les traits épais pointillés, les liaisons hydrogènes ; les traits fins pleins, les interactions d'empilement entre base et résidu de la protéine ; les traits fins pointillés, les interactions d'empilement entre bases.

3 L'adaptateur eIF4G

Cette sous-unité du complexe eIF4F assure la liaison entre les différents partenaires impliqués dans le recrutement de l'ARNm. En effet, cette protéine contient des sites de liaison d'eIF4E, PABP et eIF4A. De plus, il a pu être identifié chez les Mammifères une interaction du domaine central d'eIF4G avec eIF3 et avec l'ARN viral. Enfin, le domaine C-terminal du facteur eIF4G humain lie les kinases Mnk1 et Mnk2 ([Korneeva et al., 2000](#); [Pestova et al., 1996b](#)).

La structure tridimensionnelle d'un complexe entre les protéines de levure eIF4E liée à l'analogue m⁷GTP et eIF4G a été résolue par RMN ([Gross et al., 2003](#); [Hershey et al., 1999](#)). Lorsqu'il lie eIF4E, eIF4G s'ordonne en formant une structure hélicale de 5 segments organisés en un cercle fermé sur lui-même (**Figure 1.26**). La partie interne du cercle forme une poche hydrophobe semblable à un « bracelet » autour du domaine N-terminal d'eIF4E. Le bracelet d'un diamètre d'environ 15Å accommode une chaîne polypeptidique étendue formant le « poignet » d'un « poing » formé par les résidus N-terminaux non structurés dans la protéine libre. La surface de liaison est localisée sur la face opposée au domaine de liaison de la coiffe. Cette interaction eIF4G:eIF4E permet une stabilisation accrue de la coiffe par eIF4E. Un motif YXXXXLΦ (où X est un acide aminé quelconque et Φ un résidu aromatique) dans la séquence d'eIF4G est principalement impliqué dans la liaison d'eIF4E. Le phénomène observé chez la levure est certainement transposable au système humain compte tenu la grande conservation de séquence.

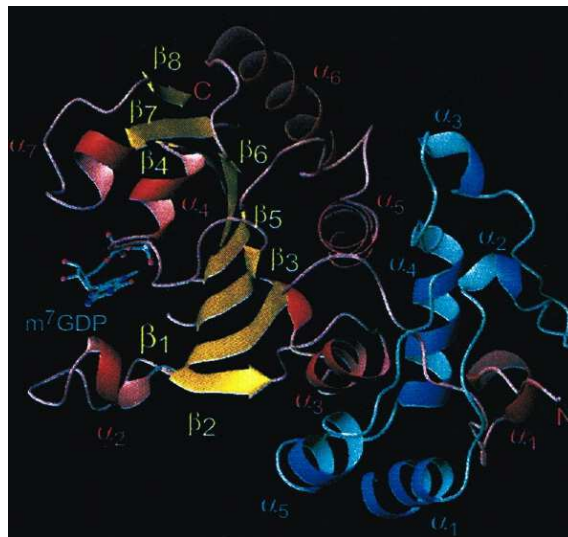


Figure 1.26 : Structure tridimensionnelle du complexe eIF4E:m⁷GTP:eIF4G(résidus 393-490) d'après ([Gross et al., 2003](#)). eIF4E est représenté en jaune et rouge et eIF4G en bleu. Les portions 1-23 d'eIF4E et 393-413 d'eIF4G n'apparaissent pas car elles sont désordonnées.

La PABP interagit avec eIF4G par les domaines RRM 1 et 2 impliqués dans la liaison de la queue poly(A) ([Kessler and Sachs, 1998](#)). Le domaine N-terminal d'eIF4G impliqué dans cette reconnaissance est cependant divergent en séquence entre les facteurs humains et de levure même si l'interaction est universellement conservée ([Imataka et al., 1998](#)). A ce jour, il n'existe pas de structure de ce complexe. Cependant, la résolution d'un complexe entre ce domaine d'eIF4G et la protéine virale NSP3 (Non Structural Protein 3) a permis de proposer des résidus de la PABP impliqués dans sa liaison à eIF4G ([Groft and Burley, 2002](#)). La protéine rotavirale NSP3 est en effet un homologue fonctionnel de la PABP à laquelle elle se substitue lors de l'infection ([Vende et al., 2000](#)). Le modèle proposé tient compte du fait que la liaison de NSP3 sur eIF4G repose essentiellement sur des contacts de Van der Waals. L'importance de ces résidus a pu être confirmée par mutagénèse ([Groft and Burley, 2002](#)). La surface d'interaction de la PABP permettant sa liaison à eIF4G serait ainsi opposée au site de liaison de la queue poly(A).

L'interaction avec la PABP stimule la formation du complexe entre eIF4E et eIF4G. Ainsi, le complexe cap:eIF4E:eIF4G:PABP:poly(A)-mRNA induit une pseudo-circularisation de l'ARNm qui pourrait provoquer l'effet synergique de la liaison des facteurs par rapprochement des partenaires ([Borman et al., 2000](#); [Otero et al., 1999](#)). Une telle circularisation a pu être visualisée par microscopie à force atomique en reconstituant *in vitro* des complexes ribonucléoprotéiques comprenant un ARNm, les facteurs eIF4E et eIF4G et la PABP ([Wells et al., 1998](#)) (**Figure 1.27**).

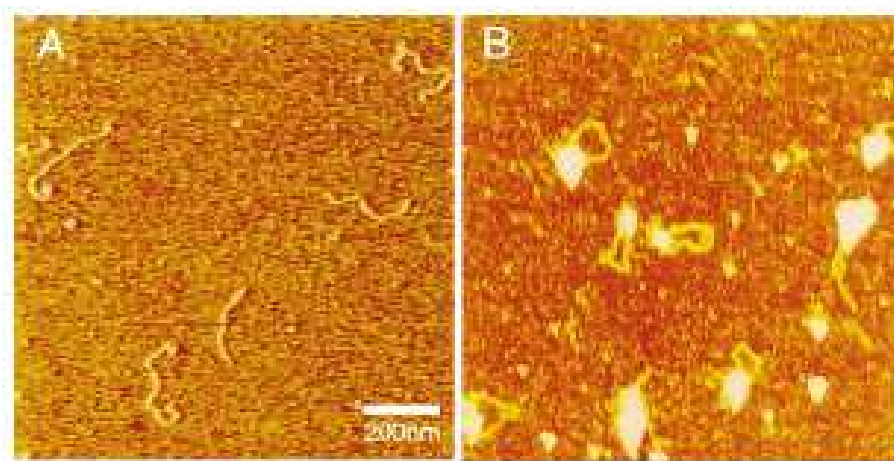


Figure 1.27 : Visualisation de complexe ARN/protéines circulaires par microscopie à force atomique d'après ([Wells et al., 1998](#)). (A) Vue de l'ARNm seul, il apparaît sous la forme de structures étendues. (B) ARNm en présence d'eIF4G, PABP et eIF4E révélant la présence de structures circulaires.

4 L'hélicase eIF4A

Le facteur eIF4G humain possède deux sites de liaison pour eIF4A alors que le facteur de levure n'en possède qu'un. Les raisons de cette différence restent obscures. eIF4A appartient à la famille des ARN-hélicases à boîte DEAD ATP-dépendante. Cette protéine possède le corps conservé des ARN-hélicases d'environ 400 résidus ainsi que des extensions de longueurs variables. eIF4A s'organise en deux domaines α/β liés par un domaine étendu non structuré conférant une grande mobilité (**Figure 1.28**). La seule structure tridimensionnelle déterminée pour ce facteur correspond à sa forme ouverte non active, sans partenaire ARN ([Caruthers et al., 2000](#)). Les séquences caractéristiques de cette famille ont pu être clairement identifiées et sa conformation active déduite de structures d'autres hélicases ([Oberer et al., 2005](#); [Sengoku et al., 2006](#)) (**Figure 1.28**). La fixation et l'hydrolyse d'une molécule d'ATP induirait un changement de conformation d'eIF4A et modifierait son affinité pour l'ARNm, ce qui pourrait transmettre l'énergie requise pour rompre les appariements au sein de l'ARNm ([Lorsch and Herschlag, 1998a, b](#)). La protéine seule a une activité très faible et son incorporation au sein du complexe eIF4F permet de l'activer ([Jaramillo et al., 1991](#); [Ray et al., 1985](#); [Rozen et al., 1990](#)). De plus, le facteur eIF4B stimulerait l'activité hélicase d'eIF4A en augmentant son affinité pour l'ARNm. Une interaction avec la PABP permet de recruter eIF4B sur le complexe ribosomal ([Fleming et al., 2003](#)). De plus, eIF4B possède un domaine qui pourrait interagir avec l'ARNr et donc stimuler le recrutement du complexe eIF4F au ribosome. Finalement, un autre facteur, eIF4H, homologue de la partie N-terminale d'eIF4B permet de stimuler l'activité hélicase d'eIF4A mais dans une moindre mesure ([Dmitriev et al., 2003](#); [Rogers et al., 2001](#)).

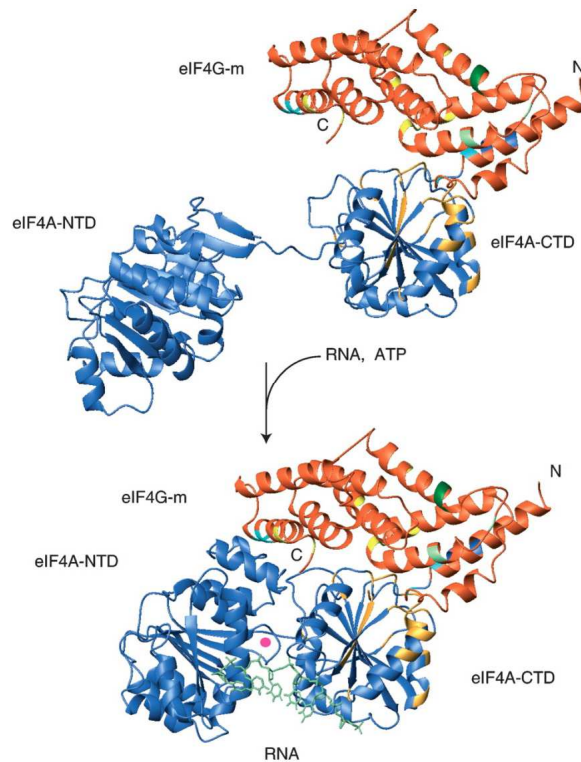


Figure 1.28 : Modèle du mécanisme d'action de la portion centrale d'eIF4G sur la configuration des domaines d'eIF4A d'après (Oberer et al., 2005). eIF4A (bleu) sous sa forme ouverte (haut) correspond à la structure cristallographique du facteur (Caruthers et al., 2000). Le domaine central d'eIF4G comportant des répétitions HEAT, représenté en orange, correspond à la structure obtenue par cristallographie (Marcotrigiano et al., 2001). Ce domaine se différencie des domaines HEAT classiques par un angle de -45° présent systématiquement entre deux motifs de répétitions successifs autour de l'axe de la superhélice. La structure ainsi créée est plus compacte. Le modèle propose que le facteur eIF4G (orange) aide à stabiliser l'hélicase sous sa forme active fermée. Les résidus d'eIF4A impliqués dans cette liaison sont indiqués en doré. Les résidus d'eIF4G importants pour maintenir l'activation d'eIF4A sont représentés en bleu et vert alors que la mutation des résidus apparaissant en doré n'a aucun effet.

Un domaine HEAT dans la portion centrale du facteur eIF4G assure la liaison d'eIF4A chez tous les Eucaryotes. La structure tridimensionnelle révèle un arrangement typique en superhélice droite (Marcotrigiano et al., 2001) (Figure 1.28). En outre, un domaine C-terminal spécifique aux mammifères présente deux domaines HEAT atypiques en tandem. Le premier domaine permet de lier un second facteur eIF4A alors que le second lie les kinases Mnk1 et Mnk2. Les domaines sont liés par une région flexible et ces domaines ne sont probablement pas en contact. La stimulation de la liaison de l'ARNm à l'hélicase lorsqu'elle se trouve au sein du complexe eIF4F serait ainsi liée à une stabilisation de la forme active (fermée) des domaines d'eIF4A une fois fixés au domaine HEAT d'eIF4G (Oberer et al., 2005; Pestova et al., 1996b) (Figure 1.28).

5 Liaison du complexe eIF4F:ARNm au complexe 43S

Le facteur eIF4G possède un site de liaison avec des sous-unités humaines d'eIF3 n'ayant pas d'homologues chez la levure. Jusqu'à présent, chez la levure, aucune interaction entre eIF4G et eIF3 n'a pu être mise en évidence. En fait, chez l'humain, le complexe eIF4F serait recruté par deux réseaux d'interactions. Le premier implique les facteurs eIF3, eIF4B et la sous-unité 40S. Le deuxième fait interagir eIF4A, eIF4B et probablement la sous-unité β d'eIF2 sur l'ARNm (Lamphear et al., 1995; Methot et al., 1996). Certaines interactions apparaissent donc redondantes et cela pourrait expliquer l'absence d'interaction entre eIF3 et eIF4G chez la levure. Le facteur eIF3 pourrait ainsi interagir indirectement avec le facteur eIF4G via ses

interactions avec eIF5 et eIF1 qui lie eIF3c ([Asano et al., 2001b](#); [He et al., 2003](#); [Yamamoto et al., 2005](#)). L'effet de la liaison d'eIF1 et eIF1A sur la conformation de la sous-unité 40S a pu être caractérisé par cryomicroscopie électronique ([Passmore et al., 2007](#)) (**Figure 1.29 (A), (B)**). Seuls, eIF1 ou eIF1A n'induisent que de faibles réarrangements de la sous-unité ribosomale. En revanche, lorsqu'eIF1 et eIF1A sont ajoutés ensemble, la tête de la sous-unité voit sa conformation modifiée, aboutissant à une ouverture plus importante du canal d'entrée de l'ARNm. En effet, il est observé la destruction du pont entre la tête et le corps de la sous-unité 40S alors qu'une nouvelle connexion est observée entre la tête et l'épaule.

Un remodelage de la sous-unité 40S sous l'effet de la liaison d'eIF1A et du MFC a également pu être visualisé par cryomicroscopie électronique ([Gilbert et al., 2007a](#)) (**Figure 1.29 (C)**). Cette étude confirme l'élargissement du canal de liaison de l'ARNm par une reconfiguration des liens entre tête, plateforme et corps. Ces modifications impliqueraient une réorganisation des parties accessibles au solvant plutôt qu'une modification de liaisons non-covalentes au sein de la sous-unité. L'ouverture de l'espace entre la tête, d'une part, et le corps et la plateforme, d'autre part, représenterait un état du ribosome actif pour le recrutement du complexe eIF4F ([Gilbert et al., 2007a](#)). Le positionnement du complexe eIF4F:mRNA:eIF4B sur le ribosome reste néanmoins encore largement à préciser.

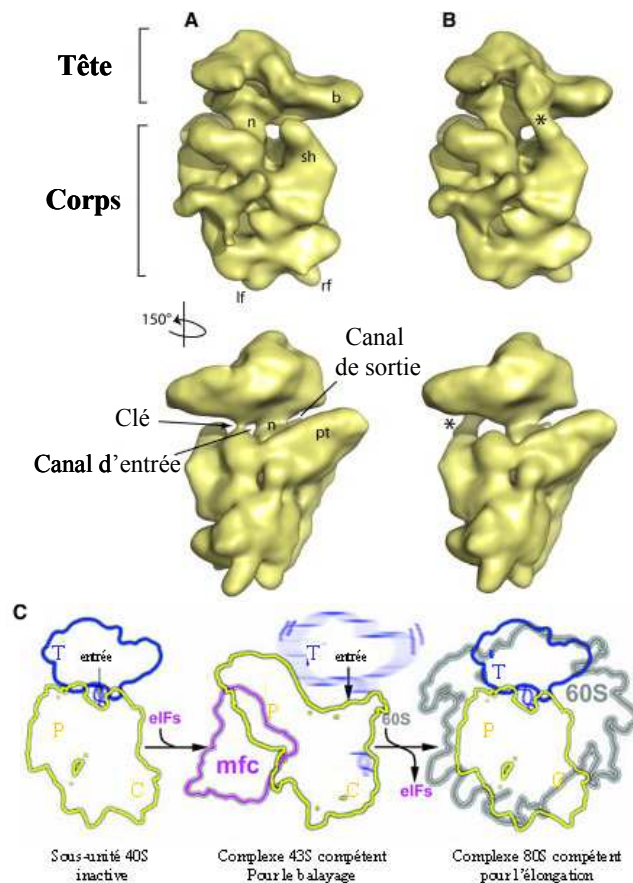


Figure 1.29 : Schéma des changements de conformations de la sous-unité 40S par la liaison des facteurs de démarrage. (A) (B) Changements de conformation de la sous-unité 40S sous l'effet de la liaison d'eIF1 et eIF1A d'après ([Passmore et al., 2007](#)). (A) Structure de la sous-unité 40S. (B) Structure du complexe 40S:eIF1:eIF1A. La tête, le bec (b), le cou (n), l'épaule (sh), le corps, la plateforme (pt), le pied droit (rf), le pied gauche (lf) et les canaux d'entrée et de sortie de l'ARNm sont indiqués. La clé qui ferme le canal de liaison de l'ARNm et la nouvelle liaison observée après fixation des facteurs eIF1 et eIF1A sont identifiées(*). (C) Modèle du réarrangement de la petite sous-unité ribosomale lors du démarrage de la traduction d'après ([Gilbert et al., 2007a](#)). La tête (T), le corps (C), la plateforme (P), le complexe multifacteurs (mfc) et le canal d'entrée de l'ARNm sont indiqués.

La structure d'eIF3 sur le ribosome obtenue par cryomicroscopie électronique permet de proposer un modèle de liaison d'eIF4G sur la sous-unité ribosomale au regard de sa liaison avec son partenaire (**Figure 1.22**). La localisation du facteur proche du site E indiquerait que le ribosome ne puisse inspecter l'ARNm qu'après un certain nombre de nucléotides. Cela concorde avec l'observation du nombre peu important d'ARNm avec des séquences 5' non codantes courtes et le peu d'efficacité de leur traduction ([Kozak, 1991](#)).

6 Le balayage de l'ARNm

Le modèle de balayage proposé en 1978 par Marilyn Kozak fournit le mécanisme le plus cohérent au regard de l'état des connaissances. Le balayage inclut de nombreux processus : le déroulement des structures secondaires de l'ARNm, le mouvement du ribosome, la discrimination d'appariements codon-anticodon non canonique et enfin la sélection d'un codon de démarrage correct.

L'observation d'un balayage ATP indépendant lorsque l'extrémité 5' non codante ne comporte pas de structures secondaires indique le caractère intrinsèque de l'habilité du ribosome à balayer l'ARNm. Un tel balayage a pu être constaté en présence des seuls facteurs eIF1, eIF1A, eIF3 et du complexe ternaire ([Pestova and Kolupaeva, 2002](#)). Les mécanismes impliquent certainement le remodelage du ribosome observé en cryomicroscopie électronique (**Figure 1.29**). Le processus requiert néanmoins la présence d'eIF4G et eIF4A lorsque des structures secondaires pourraient bloquer la progression du ribosome ([Pestova and Kolupaeva, 2002](#)). Le complexe reste donc certainement suffisamment proche du ribosome pour assurer qu'un ARNm déroulé pénètre dans le canal de liaison de l'ARNm. En outre, la stabilité du complexe entre eIF4E et la coiffe au sein d'eIF4F indiquerait sa persistance au cours du balayage. Et, la présence de ces facteurs au cours du balayage est de plus en plus suspectée ([Poyry et al., 2004](#)). Le déroulement des structures secondaires pourrait dans ce contexte être couplé au mouvement du ribosome. Si la localisation du complexe proche du site E est confirmée, le complexe pourrait « tirer » l'ARNm dans le canal de la petite sous-unité ribosomale. La reformation des structures secondaires à la sortie du canal pourrait assurer un balayage directionnel, propriété intrinsèque au ribosome, en l'empêchant de reculer et éventuellement être aussi un moteur du mouvement.

III Sélection du codon de démarrage et réponse à l'appariement

1 Sélection du codon de démarrage

Le processus de reconnaissance du codon de démarrage repose essentiellement sur la complémentarité avec l'anticodon de l'ARNt initiateur. Ainsi, le ribosome assisté des facteurs de démarrage doit pouvoir discriminer les codons AUG des codons non-AUG et déstabiliser des appariements faibles avec l'anticodon de l'ARNt initiateur. Les études biochimiques attribuent au facteur eIF1 un rôle majeur. Celui-ci peut en effet discriminer des codons de démarrage dans un environnement défavorable ou situés trop près de l'extrémité 5' ([Pestova and Kolupaeva, 2002](#)). De plus, le facteur peut dissocier des complexes ribosomaux préassemblés sur un appariement aberrant ([Lomakin et al., 2006](#); [Pestova et al., 1998](#); [Pestova and Kolupaeva, 2002](#)). Des études génétiques chez la levure sont en accord avec l'implication dans la sélection du codon de démarrage : des mutations dans une zone très conservée du facteur permettent un démarrage sur un codon UUG ([Yoon and Donahue, 1992](#)). En fait, le facteur peut réaliser son activité discriminatrice dans des systèmes hétérologues et notamment se substituer à son homologue structural bactérien IF3 ([Dallas and Noller, 2001](#); [Lomakin et al., 2003](#); [Lomakin et al., 2006](#)). Le parallèle avec IF3 indiquerait alors un contrôle plus indirect dans la sélection, en induisant notamment des changements conformationnels du ribosome (**Figure 1.29**). De plus, des études proposent également un rôle d'eIF1 dans la sélection de l'ARNt initiateur au site P par inspection des paires G:C de la tige anticodon, assistée par les résidus de l'ARNr 18S correspondant aux résidus G1338 et A1339 de l'ARNr 16S d'*E. coli*

([Lancaster and Noller, 2005](#); [Lomakin et al., 2006](#)). Des études portant sur les cinétiques de la sélection du codon initiateur indiquent que le facteur eIF1 stimule un changement de conformation de la petite sous-unité ribosomale dont la structure est alors dite « ouverte », c'est-à-dire compétente pour le balayage et assurant la fidélité du processus. Suivant l'appariement codon-anticodon, le complexe 48S subirait une nouvelle réorganisation, adoptant une conformation dite « fermée », fixant la position de l'ARNt et de l'ARNm et donc le cadre de lecture ([Pestova and Hellen, 2000](#); [Pestova and Kolupaeva, 2002](#)) (**Figure 1.29**). Ce modèle a pu être confirmé chez la levure en y ajoutant l'intervention conjointe du facteur de démarrage eIF5 ([Maag et al., 2006](#)). Alors que le facteur eIF1A stabilise la formation du complexe sous sa forme ouverte, le facteur eIF1 exerce une action sur la cinétique de formation du complexe mais pas sur sa stabilité ([Passmore et al., 2007](#)).

L'intervention de la séquence consensus GCC(A/G)CCAUGG (la position dite +1 correspond à l'adénine du codon initiateur ; les bases en gras indiquent les positions les plus importantes) identifiée par Kozak en 1987 est longtemps restée énigmatique. En effet, aucun segment de l'ARNr 18S ne semble contenir une séquence qui en serait complémentaire, à la différence du prépositionnement offert par la séquence bactérienne de SD. Néanmoins, des résultats récents indiquent que la base G en position +4 interagit spécifiquement avec les bases A1818-A1819 et que la base G en position -3 interagit avec eIF2 α ([Pisarev et al., 2006](#)). L'implication d'eIF2 dans le processus de reconnaissance du codon de démarrage est en accord avec une étude qui a identifié eIF5 et eIF2 comme des acteurs de ce mécanisme ([Huang et al., 1997](#)).

Les facteurs eIF1A, eIF4 et eIF3 ont également été impliqués dans l'étape de sélection du codon de démarrage, certainement par le réseau d'interaction les liant à eIF1, eIF2 et eIF5 ([Fekete et al., 2005](#); [He et al., 2003](#); [Nielsen et al., 2004](#); [Valasek et al., 2004](#)). L'implication d'eIF1A n'est pas surprenante au regard de l'implication de son homologue bactérien dans ce même processus. La reconstitution *in vitro* de complexe de prédémarrage a de plus pu mettre en évidence une interaction critique entre les facteurs eIF1A et eIF5, AUG-dépendante, qui influencerait l'équilibre entre les conformations « ouverte » et « fermée » du complexe ([Maag et al., 2006](#)).

2 Hydrolyse du GTP

Suite à l'appariement codon-anticodon, le complexe ribosomal dans sa conformation « fermée » doit s'associer à la grande sous-unité ribosomale. Or les facteurs eIF1, 1A, 2 et 3 sont localisés sur des sites de liaison de la sous-unité 60S, ils doivent donc être déplacés de cette surface. Cette étape nécessite l'hydrolyse du GTP associé à eIF2, qui intervient avec l'aide du facteur eIF5, la GAP spécifique à eIF2.

Le facteur eIF5 se lie à eIF2 par des interactions avec ses sous-unités β et γ ([Alone and Dever, 2006](#); [Das et al., 1997](#)). L'activation de l'activité GTPasique du facteur eIF2 par le facteur eIF5 est optimale lorsque ce dernier est lié au complexe ternaire associé à la sous-unité 40S ([Algire et al., 2005](#)). Le mode d'action d'eIF5 en tant que GAP demeure incertain. Le facteur pourrait agir comme une GAP « classique » à savoir en stabilisant l'état de transition de l'hydrolyse du GTP. Ceci serait soutenu par l'observation que seul le domaine N-terminal contenant le site de liaison à eIF2 γ est suffisant pour l'activité d'eIF5 ([Alone and Dever, 2006](#)). L'interaction entre le domaine C-terminal d'eIF5 et eIF2 β stimulerait le recrutement d'eIF5 sur le complexe 43S. Une autre hypothèse est qu'eIF5 interagit avec eIF2 β afin de lever la répression de cette sous-unité sur l'activité GTPasique d'eIF2 γ . Les effets de mutations dans la sous-unité β d'eIF2 qui diminuent son affinité pour γ et conduisent à une activité GTPasique intrinsèque eIF5-indépendante seraient cohérents avec ce mode d'action d'eIF5 ([Hashimoto et al., 2002](#); [Huang et al., 1997](#)).

Le modèle précédent de la réponse à un appariement correct codon-anticodon stipulait que l'hydrolyse du GTP par eIF5 conditionne une modification irréversible du complexe 48S. Ainsi, l'activité GAP d'eIF5 interviendrait lors de la reconnaissance de l'appariement codon-anticodon au site P ([Unbehaun et al., 2004](#)). Cependant, au sein d'un complexe contenant les eIFs 1, 1A, 2 et 5, il apparaît que l'éjection du phosphate, et non l'hydrolyse du GTP est l'étape déclenchée par l'appariement ([Algire et al., 2005](#)). Dans ce contexte, il faut noter que la structure du facteur hétérotrimérique $\alpha_3\beta\gamma$ du facteur aIF2 obtenue récemment au laboratoire fournit des éléments intéressants ([Yatime et al., 2007](#)). Le facteur aIF2 est, dans cette structure, sous sa forme liée au GDP. L'inspection de la structure révèle néanmoins une conformation des « switches » ne correspondant pas à une forme inactive. En effet, des résidus des régions « switches » pointent vers une densité électronique qui peut être attribuée à un groupement phosphate. Les conformations des « switches » sont particulières, le « switch » 1 est dans une conformation nouvelle et « switch » 2 est en position « on ». Ainsi, cette structure correspondrait au facteur ayant subit hydrolysé le GTP sans avoir éjecté le phosphate. La liaison de l'ARNt pourrait toujours avoir lieu. Cette structure illustrerait les résultats d'Algire obtenus pour le système eucaryotique ([Algire et al., 2005](#)).

L'hydrolyse du GTP pourrait ainsi avoir lieu avant la reconnaissance de l'appariement, par exemple au cours du balayage (**Figure 1.30**). Le modèle actuel considère que l'irréversibilité de l'étape est conditionnée par l'éjection du phosphate d'eIF2.

3 Ejection du phosphate

La connaissance du système archéen conduit à certaines remarques. En effet, il n'est pas identifié à ce jour de GAP spécifique à aIF2, la sous-unité β du facteur ne possède pas le domaine d'interaction avec la GAP mais une activité GTPasique intrinsèque au facteur aIF2 n'a pu être démontrée ([Londei, 2005](#)). De plus, chez les Archées, l'ARNm est prépositionné, il n'y a pas de balayage. Ceci renforce l'idée que l'activité d'eIF5 soit intimement liée au processus de balayage. Ainsi, ce facteur permettrait d'assurer l'intégrité du complexe ternaire tout au long du balayage de l'ARNm afin d'éviter une éjection prématurée du phosphate.

Chez les Eucaryotes, le facteur eIF1 inhibe l'éjection du phosphate. Son action est stimulée par l'ensemble des facteurs de démarrage et surtout par eIF3 et eIF1A ([Majumdar and Maitra, 2005](#)). De plus, il est proposé qu'eIF1 empêche la formation du complexe de prédémarrage sous sa forme fermée ([Pisarev et al., 2006](#)). La répression apportée par la présence d'eIF1 serait alors levée lors de l'établissement de l'appariement codon-anticodon, couplée aux changements conformationnels aboutissant à la forme fermée du complexe ([Algire et al., 2005](#); [Passmore et al., 2007](#)) (**Figure 1.30**). Une étude récente analysant les effets de mutations qui diminuent l'affinité d'eIF1 pour le ribosome confirme que cette interaction est critique pour le choix d'un codon de démarrage approprié ([Cheung et al., 2007](#)).

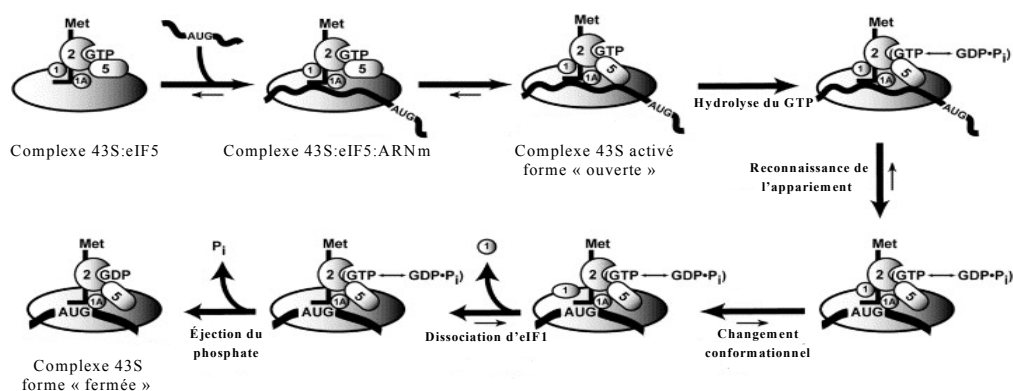


Figure 1.30 : Modèle de reconnaissance de l'appariement codon-anticodon d'après ([Algire et al., 2005](#)).

Peu d'éléments permettent de proposer les réorganisations du complexe de prédémarrage conduisant à l'éjection du phosphate. Il apparaît néanmoins que, dans ce contexte, le rôle antagoniste d'eIF1 serait plus indirect ([Algire and Lorsch, 2006](#)). Suivant l'hydrolyse du GTP, le facteur eIF5 serait maintenu dans une configuration où il bloque le phosphate au sein du domaine G d'eIF2. Suivant l'appariement codon-anticodon, la transition vers la forme « fermée » du complexe nécessiterait une modification de la position d'eIF1 ou son éjection du complexe 43S ([Maag et al., 2005](#)). Ceci permettrait de lever la répression apportée par eIF1, aboutissant à l'éjection du phosphate et à la réorganisation du complexe 43S sous sa forme « fermée » ([Algire et al., 2005](#); [Passmore et al., 2007](#)). Le facteur eIF2, lié au GDP perd son affinité pour l'ARNt et est alors relâché du ribosome probablement avec eIF5 ([Unbehaun et al., 2004](#)).

4 Régénération du facteur eIF2 : eIF2B

L'affinité du GDP pour eIF2 est 100 fois plus élevée que celle du GTP, ce qui conduit à une vitesse de dissociation du GDP très lente ([Kapp and Lorsch, 2004a](#)). La régénération de la forme active eIF2:GTP est réalisée par eIF2B. eIF2B contient 5 sous-unités : α , β , γ , δ et ϵ . Toutes les sous-unités, exceptée α , sont des protéines essentielles. Néanmoins, la sous-unité ϵ seule possède une activité GEF, localisée dans les 200 résidus C-terminaux ([Gomez et al., 2002](#); [Gomez and Pavitt, 2000](#); [Pavitt et al., 1998](#)). Ce segment contient deux boîtes AA, à l'instar d'eIF5, requises pour sa liaison au facteur eIF2 (**Figure 1.31**). La délétion ou la mutation de ces boîtes détruit la liaison de la sous-unité à eIF2 et sa fonction GEF et réduit l'activité du complexe eIF2B ([Asano et al., 1999](#)). L'extension N-terminale de la sous-unité β du facteur eIF2 est impliquée dans la liaison de la sous-unité ϵ d'eIF2B par ses boîtes Lysine (« boîtes K ») ([Asano et al., 1999](#); [Gomez et al., 2002](#); [Gomez and Pavitt, 2000](#)). La région précédant les boîtes AA du domaine C-terminal d'eIF2 ϵ est impliquée dans l'activité GEF mais pas dans la liaison à eIF2. La structure tridimensionnelle du fragment catalytique C-terminal révèle un domaine de répétitions HEAT comprenant les boîtes AA ([Boesen et al., 2004](#)). Ce fragment est fortement homologue structurellement aux motifs HEAT identifiés dans les structures des fragments C-terminaux d'eIF5 et d'eIF4G ([Bellolell et al., 2006](#); [Wei et al., 2006](#)).

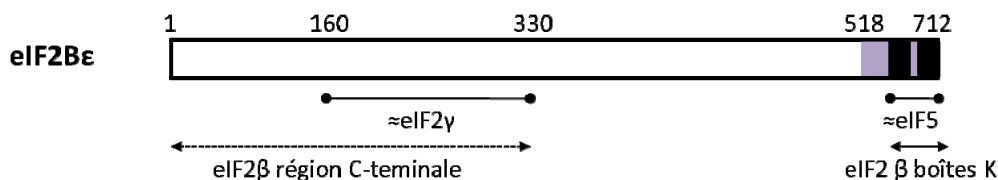


Figure 1.31 : Domaines importants du facteur eIF2B ϵ (numérotation de *S. cerevisiae*). Les doubles flèches indiquent les domaines de liaison et les régions d'eIF2 β concernées sont indiquées. Les parties homologues (\approx) à eIF5 et eIF2 γ sont représentées. Le domaine catalytique GEF apparaît en violet et les boîtes AA en noir.

La délétion des boîtes AA ne détruit cependant pas complètement l'activité GEF du facteur. Des interactions additionnelles doivent donc permettre la liaison à eIF2. La mutation d'un motif conservé de la sous-unité ϵ n'altère pas l'activité de la sous-unité seule mais réduit considérablement celle du complexe pentamérique, sans modifier l'intégrité du complexe ou sa liaison à eIF2 ([Gomez and Pavitt, 2000](#)). De plus, les sous-unités α , β et δ forment un sous-complexe capable de se lier à eIF2. Tout ceci indique un rôle stimulateur accompli par les autres sous-unités du complexe. Les mécanismes moléculaires assurant la fonction GEF restent à préciser.

5 L'association des sous-unités ribosomales : eIF5B

Après le départ d'eIF2 et d'eIF5, l'éjection des facteurs restant sur la petite sous-unité et l'association des sous-unités ribosomales est alors assurée par le facteur eIF5B, l'orthologue du facteur bactérien IF2. Le facteur possède une activité GTPasique stimulée en présence des sous-unités ribosomales 40S et 60S, à l'instar du facteur IF2 ([Merrick et al., 1975](#); [Pestova et al., 2000](#); [Shin et al., 2002](#)). Le facteur eIF5B n'est pas essentiel chez la levure mais la délétion de son gène (*FUN12*) produit un phénotype de croissance très lente. Le facteur archéen aIF5B possède une structure en quatre domaines ressemblant à un calice ([Roll-Mecak et al., 2000](#)) (**Figure 1.32**). Le domaine I du facteur correspond au domaine G. Le domaine II forme un tonneau β . Le domaine III est du type α/β . Les domaines I, II et III forment la coupe d'un calice connectée au domaine IV par une longue hélice α de 40Å. Le domaine IV se termine par un sous-domaine hélical. Il est à noter que le facteur aIF5B est dénué d'une extension N-terminale de longueur variable chez les Eucaryotes. Néanmoins, cette extension n'est pas essentielle à la fonction du facteur. En effet, le facteur archéen peut se substituer au facteur de levure *in vivo* et *in vitro* ([Lee et al., 1999](#)).

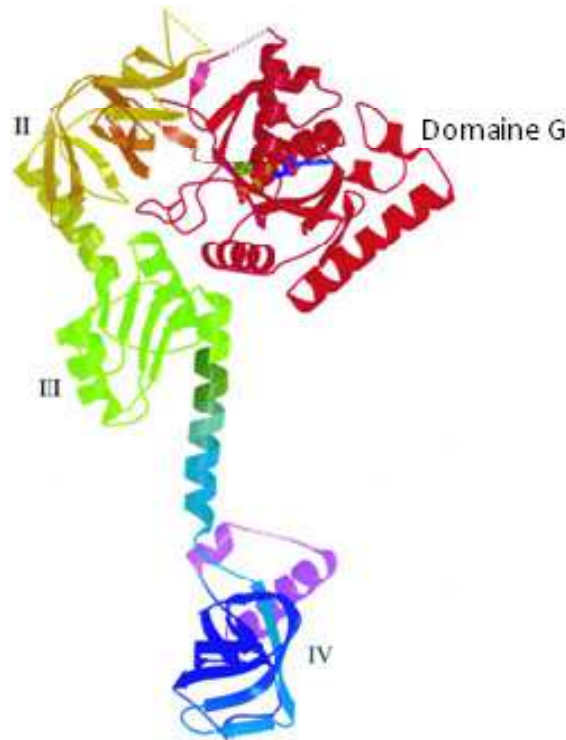


Figure 1.32 : Structure du facteur aIF5B de *Methanobacterium thermoautotrophicum* d'après ([Roll-Mecak et al., 2000](#)). Le code couleur est comme suit : domaine G, rouge ; domaine II, jaune ; domaine III, vert, domaine IV, bleu ; hélices C-terminales, violet.

Le domaine I possède les régions « switches » typiques des GTPases. Le « switch » 2 est impliqué dans des liaisons avec le cofacteur GTP-Mg^{2+} mais pas avec le GDP. Ce changement de conformation du « switch » 2 s'accompagne d'une rotation du domaine II amplifiée et transmise au domaine IV via le domaine III. Le domaine IV se voit déplacé d'environ 5Å par ce mécanisme semblable à un bras de levier ([Roll-Mecak et al., 2000](#)) (**Figure 1.33**). La reconstitution d'un système de démarrage humain *in vitro* a montré que l'hydrolyse du GTP par eIF5B n'est pas requise pour l'association de la sous-unité 60S mais pour stimuler un changement de conformation du ribosome 80S en un complexe activé compétent pour la traduction qui provoque l'éjection du facteur ([Pestova et al., 2000](#)). Cet effet est confirmé par l'analyse de mutations dans les régions switches qui détruisent l'activité GTPase ribosome-dépendante du facteur mais pas l'association des sous-unités ribosomales et le facteur muté

reste associé au ribosome 80S inactif ([Lee et al., 2002](#); [Shin et al., 2002](#)). De plus, il a pu être remarqué qu'une mutation altérant l'activité GTPase du facteur peut-être contrebalancée par une autre mutation qui réduit l'affinité d'eIF5B de levure pour le ribosome ([Shin et al., 2002](#)). Le fonctionnement d'eIF5B serait donc similaire à celui d'IF2 ([Antoun et al., 2003](#)). Il est à noter qu'un schéma cinétique précis a été décrit pour le système bactérien, caractérisant notamment la chronologie de l'hydrolyse du GTP et le rôle d'IF2 ([Grigoriadou et al., 2007a](#)). Malheureusement, aucune étude comparable dans le système eucaryotique ou archéen ne permet de comparer le mode d'action précis d'eIF5B et celui d'IF2.

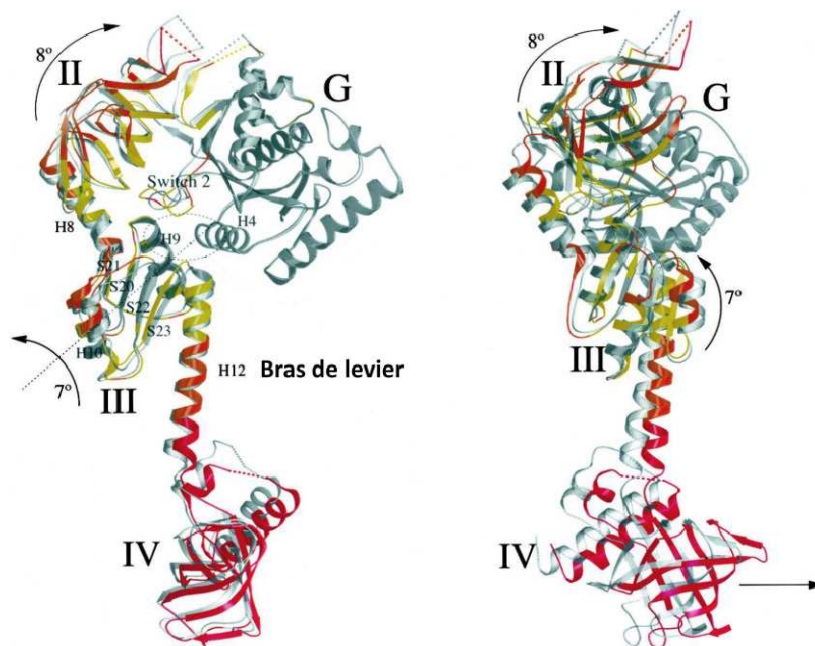


Figure 1.33 : Visualisation du mouvement des domaines d'eIF5B induit par la liaison du GTP-Mg²⁺ d'après ([Roll-Mecak et al., 2000](#)). Les vues correspondent à une rotation de 90° selon un axe vertical. Les domaines G d'eIF5B:GDP (gris) et d'eIF5B:GTP (colorée suivant l'amplitude du mouvement des résidus : <1 Å, gris; 1–2 Å, jaune; 2–3 Å, orange clair; 3–4 Å, orange foncé; >4 Å, rouge) ont été alignés. La numérotation des domaines est indiquée et leurs mouvement relatifs indiqués par des flèches.

Il a été observé *in vitro* et *in vivo* une association entre les facteurs de démarrage eIF1A et eIF5B. Une interaction entre les facteurs IF1 et IF2, les homologues bactériens d'eIF1A et eIF5B, a également été mise en évidence sur le ribosome ([Boileau et al., 1983](#)). L'interaction reportée pour les facteurs eucaryotiques est néanmoins totalement spécifique à ces organismes puisqu'elle implique l'extension C-terminale d'eIF1A spécifique aux Eucaryotes ([Olsen et al., 2003](#)). Cependant, une autre interaction entre eIF1A et eIF5B, du type IF1:IF2, pourrait également avoir lieu sur le ribosome ([Marintchev et al., 2003](#)).

Une analyse de titration par RMN révèle que les résidus d'eIF1A impliqués dans la liaison spécifique aux Eucaryotes interagissent avec une région hydrophobe formée par les hélices C-terminales d'eIF5B ([Marintchev et al., 2003](#)). L'association nécessite les parties C-terminales des deux protéines (24 résidus C-terminaux pour eIF1A et 153 résidus C-terminaux pour eIF5B) ([Choi et al., 2000](#); [Olsen et al., 2003](#)). L'effet de mutations au sein des sites de l'interaction spécifique aux Eucaryotes révèle que l'association est nécessaire *in vivo* ([Choi et al., 2000](#); [Olsen et al., 2003](#)). Cette interaction eIF5B:eIF1A pourrait fournir une base pour le recrutement d'eIF5B au ribosome. L'interaction avec eIF1A stimule l'association des sous-unités ribosomales et l'hydrolyse du GTP lié à eIF5B *in vitro* ([Acker et al., 2006](#)). Enfin, la surexpression d'eIF1A dans une souche de levure dont le gène *FUN12* a été muté de sorte à produire eIF5B dénuée de son domaine C-terminal exacerbe le phénotype de croissance lente.

Il est proposé que l'éjection d'eIF1A du ribosome soit dépendante de son interaction avec eIF5B. Dans cette souche, l'impossibilité d'éjecter eIF1A de son site de fixation sur le ribosome au site A empêcherait la première étape d'élongation (Choi et al., 2000).

Compte tenu de la position supposée d'eIF1A sur le ribosome, le domaine C-terminal d'eIF5B serait localisé au niveau du site A, proche de la tige acceptrice de l'ARNt au site P. Cette configuration serait cohérente avec la position observée pour IF2 sur la sous-unité 30S bactérienne par cryomicroscopie électronique (Allen et al., 2005; Myasnikov et al., 2005). De plus, l'étude de Myasnikov et al. permet de visualiser les changements conformationnels du facteur suivant l'hydrolyse du GTP. Il est décrit une rotation du domaine II par rapport au domaine G induisant le retrait du domaine IV du site A et, consécutivement, la perte de l'interaction avec la tige acceptrice de l'ARNt au site P ainsi qu'une diminution des interactions avec le centre GTPase de la sous-unité 60S et le centre de décodage (Myasnikov et al., 2005) (Figure 1.34). Les changements conformationnels qui ont été observés en cryomicroscopie électronique sont cohérents avec la mobilité du domaine IV par rapport aux domaines I-II-III, mise en évidence dans les structures cristallographiques du facteur archéen (Roll-Mecak et al., 2000).

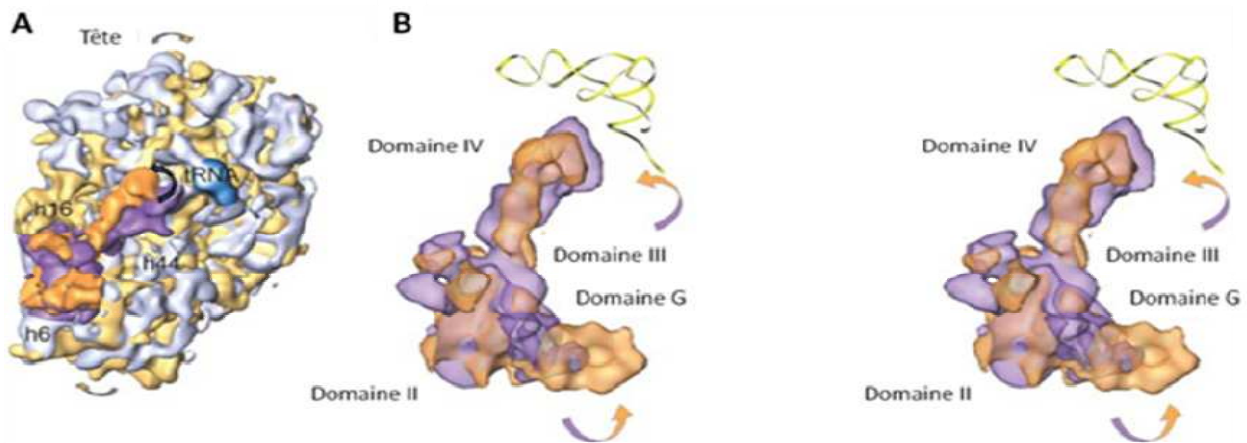


Figure 1.34 : Comparaison des états liés au GTP et au GDP d'IF2, illustrant le changement conformationnel et le repositionnement du facteur sur le ribosome induit par l'hydrolyse du GTP, d'après (Myasnikov et al., 2005). (A) Superposition du complexe IF2:GMPPCP:30S (densité du ribosome : bleu clair ; densité d'IF2 : violet) et du complexe IF2:GDP:30S (du ribosome : jaune ; densité d'IF2 : orange). La tête de la sous-unité 30S est indiquée et l'ARNt apparaît en bleu foncé. (B) Représentation stéréographique d'IF2 dans les deux états superposés. L'ARNt apparaît en jaune. Suivant la transition du GTP au GDP, IF2 subit une rotation suivant son grand axe (B) et également suivant un axe perpendiculaire (A, flèche noire). La transition s'accompagne d'une faible rotation ($\approx 5^\circ$) de la sous-unité 30S au regard de la sous-unité 50S (A, flèches).

IV Régulation & modes alternatifs du démarrage de la traduction

Le démarrage étant l'étape limitante de la traduction, de nombreux mécanismes de régulation interviennent à ce niveau. Notamment, la description de quelques mécanismes montrera l'importance des sites de phosphorylation des facteurs de démarrage ou de leurs partenaires dont la caractérisation a depuis longtemps été entreprise ([Benne et al., 1978](#)).

1 Régulation de la liaison du Met-ARNt_i^{Met} : phosphorylation d'eIF2 α

La sous-unité α de l'hétérotrimère eIF2 joue un rôle critique dans la régulation de son activité. Le domaine OB-fold de la sous-unité contient en effet la Ser51 strictement conservée chez les Eucaryotes. Les résidus avoisinant ce résidu ont un rôle important dans la liaison du facteur eIF2B ([Dey et al., 2005](#); [Krishnamoorthy et al., 2001](#); [Vazquez de Aldana et al., 1993](#)). La phosphorylation de la Ser51 n'altère pas la liaison du Met-ARNt_i^{Met} mais convertit eIF2 en un mauvais substrat pour eIF2B. En fait, la modification augmente l'affinité d'eIF2 pour eIF2B ([Pavitt et al., 1998](#)). Le facteur eIF2(α P):GDP pourrait ainsi constituer un inhibiteur compétitif ([Dever et al., 1995](#); [Rowlands et al., 1988](#)). L'effet de la phosphorylation est renforcé par la proportion plus importante d'eIF2 par rapport à eIF2B dans la cellule. Le modèle est en accord avec le fait que la phosphorylation puisse être compensée par la surproduction des sous-unités d'eIF2B impliquées dans cette régulation (α , β et δ) ([Yang and Hinnebusch, 1996](#)).

Quatre kinases ont été identifiées pour la phosphorylation sélective de la Ser51 d'eIF2 α (HR1, PKR, PERK et GCN2). Ces kinases sont activées en fonction de différents stress. Nous nous intéresserons plus particulièrement à GCN2, protéine présente chez tous les Eucaryotes, premièrement identifiée comme un stimulateur de la traduction de l'ARNm *GCN4* chez la levure. La phosphorylation par GCN2 est activée en réponse à un manque d'acides aminés ([Dever and Hinnebusch, 2005](#)). GCN2 contient un domaine fortement homologue à l'Histidyl ARNt-synthétase. Ce segment lie préférentiellement les ARNt déacylés par rapport aux ARNt aminoacylés. Cela fournit une voie d'activation de la kinase par identification d'un excès d'ARNt ne portant pas d'acide aminé ([Dong et al., 2000](#)). Ainsi, chez la levure, la phosphorylation d'eIF2 α catalysée par GCN2 résulte en la stimulation de l'expression d'un ARNm spécifique, celui de *GCN4*. Cette situation semble conservée chez tous les Eucaryotes. En effet, la diminution de la disponibilité en eIF2 actif a un effet négatif sur la traduction de la majorité des ARNm mais elle peut aussi, paradoxalement, induire la traduction d'ARNm spécifiques codant pour des protéines de réponse aux stress.

La protéine GCN4 est un activateur de la transcription des gènes impliqués dans la biosynthèse des acides aminés. Dans sa région 5' non codante, l'ARNm de *GCN4* contient 4 petits cadres de lectures ouverts alors appelés uORF (upstream Open Reading Frame) (**Figure 1.35 (A) construction a**). Dans des conditions normales, la traduction des uORF limite le flux de ribosome pouvant atteindre le codon de démarrage de *GCN4* et la protéine GCN4 est produite à un niveau basal. En revanche, lorsque la Ser51 est phosphorylée, la protéine GCN4 est surexprimée. En fait, la présence de petits ORF en amont du gène à traduire conduit à deux alternatives au démarrage canonique : le « leaky scanning » et le redémarrage ([Kozak, 2002](#)). Le « leaky scanning » (ou poursuite de balayage) correspond à un ARNm contenant un codon de démarrage dans un contexte défavorable, le ribosome passe alors le premier uORF sans y être associé et balaye l'ARNm jusqu'au second uORF. Le second mécanisme proposé, le redémarrage, intervient lorsque le ribosome est associé sur un ARNm comportant plusieurs ORF, le ribosome peut continuer à balayer l'ARNm à condition que l'ORF traduit soit suffisamment court. Dans le cas de *GCN4*, la délétion des 4 uORF lève la répression alors que la présence uniquement de l'uORF4 inhibe toute expression du gène (**Figure 1.35 (A) construction b et c**). En revanche, lorsque seul le premier uORF reste, on n'observe qu'une faible inhibition de l'expression (**Figure 1.35 (A) construction d**). Les effets différents de ces uORF sont certainement liés aux séquences nucléotidiques qui les encadrent. Des séquences riches en A+U avoisinant le codon stop d'uORF1 favorisent la continuation du balayage, et

donc le redémarrage, alors que les séquences riches en G+C proches du codon stop d'uORF4 stimulent la dissociation du ribosome après sa traduction.

Les uORF2 et 3 ayant un impact mineur sur le processus, nous nous limiterons à un ARNm modèle ne comportant que les uORF 1 et 4 pour l'expliquer (**Figure 1.35 (A) construction e**). Le mécanisme propose que dans tous les cas l'uORF1 est effectivement traduit. Lorsqu'eIF2 n'est pas phosphorylé, le ribosome qui continue le balayage après la traduction d'uORF1 acquiert un nouveau complexe ternaire. Le redémarrage peut alors avoir lieu sur uORF4 suite à quoi il se dissocie de l'ARNm (**Figure 1.31 (B) situation f**). En revanche, lorsqu'eIF2 est phosphorylé, l'effet majeur résulte en une disponibilité décruée du complexe ternaire. Ainsi, après avoir traduit uORF1, le ribosome balaye plus longuement l'ARNm avant d'acquérir un nouveau complexe ternaire et passe uORF4 sans pouvoir redémarrer sur ce codon de démarrage. C'est *GCN4* qui est alors effectivement traduit ([Dever et al., 1995](#)) (**Figure 1.35 (B) construction g**).

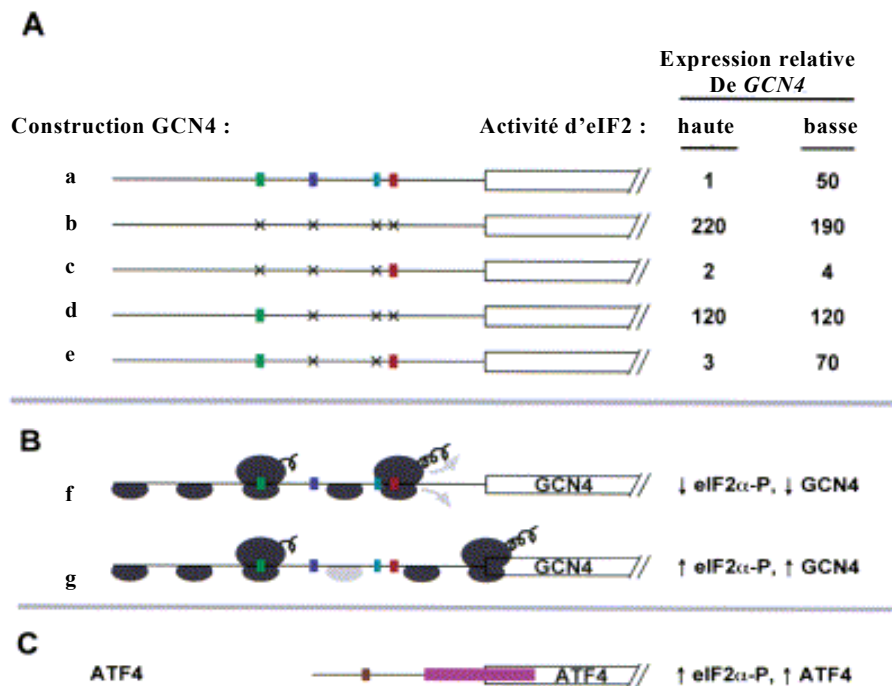


Figure 1.35 : Régulation de la traduction de *GCN4* et *ATF4* d'après ([Dever, 2002](#)). (A) Mutations des uORF de *GCN4* qui altèrent son expression. Les uORF sont repérés par des rectangles. Lorsqu'ils apparaissent sous forme de croix, le codon de démarrage a été muté. Les niveaux d'expression du gène rapporteur *GCN4-LacZ* ont été mesurés dans deux souches. La première permet une forte activité d'eIF2 alors que la seconde produit une forme mutée d'eIF2B mimant la phosphorylation d'eIF2 α et donc une faible activité d'eIF2. (B) (C) Modèle de la régulation de la traduction de *GCN4* et *ATF4*, respectivement.

Les mécanismes de la régulation de l'expression de *GCN4* sont bien compris, l'identification d'un ARNm comparable chez les mammifères : *ATF4* confirme l'universalité du phénomène ([Vattem and Wek, 2004](#)). Ce gène code également pour un activateur de la transcription en réponse à un défaut d'acides aminés. L'ARNm contient alors un premier uORF suivi d'un deuxième recouvrant le gène *ATF4* (**Figure 1.35 (C)**). De même que précédemment, un délai avant d'acquérir un nouveau complexe ternaire est nécessaire afin de passer uORF2 sans traduction et d'ainsi atteindre le gène *ATF4*. Néanmoins, il convient de préciser que ce mécanisme parfaitement caractérisé chez la levure reste un cas exceptionnel et moins de 1% de ses ARNm semblent contenir des uORF ([Zhang and Dietrich, 2005](#)). En revanche, une analyse des génomes humain et de rongeur révèle qu'environ 35% d'ARNm contenant au moins 1 uORF ([Iacono et al., 2005](#)). Bien que cela n'indique pas si ces uORF sont fonctionnels, ce phénomène de régulation pourrait être plus répandu chez les vertébrés.

2 Régulation de la liaison de l'ARNm : phosphorylation d'eIF4E et 4E-BP

Chez les mammifères, le facteur eIF4E présente une sérine universellement conservée (Ser209). Celle-ci est sélectivement phosphorylée par les kinases Mnk1 et Mnk2 liées au domaine HEAT C-terminal de la sous-unité eIF4G du complexe eIF4F ([Bellsollell et al., 2006](#); [Pyronnet et al., 1999](#)). Les conséquences de cette phosphorylation restent controversées ([Scheper and Proud, 2002](#)). Dans un premier temps, un rôle stimulateur par amélioration de l'affinité pour la coiffe a été proposé. Plus précisément, la phosphorylation permettrait de former un pont salin avec la Lys159, située au-dessus du canal liant l'ARNm, clampant ainsi l'ARNm et sa coiffe ([Marcotrigiano et al., 1997](#); [Minich et al., 1994](#)). Ces idées ont récemment été contestées, suggérant, au contraire, que la phosphorylation induirait une diminution de l'affinité pour la coiffe ([Scheper et al., 2002](#); [Zuberek et al., 2003](#)). De plus, la structure d'eIF4E lié à un analogue de la coiffe indique une distance entre les résidus Ser209 et Lys159 trop importante (19Å) pour permettre la formation d'un pont salin. La diminution de l'affinité pourrait alors être induite par des forces de répulsions électrostatiques entre le phosphate de la sérine et le phosphate de l'adénine de l'analogue de la coiffe m⁷GpppA ([Tomoo et al., 2002](#)). Malgré tout, la distance entre ces groupements de 7Å paraît trop importante pour justifier un tel effet. Une simulation informatique de l'effet de la phosphorylation prédit cependant un rétrécissement de la distance en Ser209-P et Lys159 conduisant à une poche de fixation de l'ARNm plus étroite ([Tomoo et al., 2003](#)). Cette diminution d'affinité conduirait à l'éjection d'eIF4E de la coiffe et serait donc favorable à l'entrée en balayage ([Scheper and Proud, 2002](#)).

En marge de cette phosphorylation, l'activité d'eIF4E est régulée par une classe de 3 protéines 4E-BP (4E Binding Protein) chez les Mammifères. Ces protéines inhibent la synthèse protéique *in vivo* et *in vitro*. La liaison de ces protéines sur eIF4E n'altère pas son affinité pour la coiffe ([Pause et al., 1994](#)). En revanche, en leur présence, la liaison d'eIF4E à eIF4G et donc l'assemblage d'un complexe eIF4F n'est plus possible ([Haghighat et al., 1995](#)). La famille des 4E-BP contient un motif similaire à celui que contient eIF4G, responsable de la liaison à eIF4E ([Mader et al., 1995](#)). Chez la levure, la protéine (p20) de levure possède le motif caractéristique impliqué dans liaison à eIF4E et est homologue fonctionnellement aux 4E-BP ([Altmann et al., 1997](#)). L'activité de ces protéines est fortement régulée par l'état de phosphorylation de résidus sérine/thréonine spécifiques ([Pause et al., 1994](#)). En effet, les formes hypophosphorylées de ces protéines interagissent fortement avec eIF4E alors que la phosphorylation d'un certain nombre de ces résidus réduit considérablement cette liaison. L'effet semble résulter de la phosphorylation de 5 à 6 sites dans un ordre donné ([Gingras et al., 2001](#)). L'hyperphosphorylation des 4E-BP est favorisée par de nombreux stimuli et notamment les hormones et les facteurs de croissance alors que leur hypophosphorylation intervient sous certaines conditions de stress. La région identifiée pour la liaison des 4E-BP est cohérente avec la surface d'interaction avec eIF4G ([Matsuo et al., 1997](#)). De plus, cette structure d'eIF4E avec un peptide mimant une 4E-BP indique la présence de régions acides sur eIF4E encadrant le peptide. Ainsi, la phosphorylation de ces protéines pourrait induire des répulsions électrostatiques avec ces résidus acides, diminuant leurs affinités pour eIF4E.

3 Régulation de la PABP

Le mode d'action et de régulation de cette protéine est encore mal caractérisé vu son implication dans de nombreux processus biologiques et notamment au niveau de la transcription et de la dégradation des ARNm. Au cours du démarrage, deux partenaires additionnels de la PABP permettent de réguler son activité : Paip1 et Paip2 (Paip : PABP Interacting Protein) ([Craig et al., 1998](#); [Khaleghpour et al., 2001](#)) (**Figure 1.36 (A)**). Paip1 stimule la traduction *in vivo* ([Craig et al., 1998](#)). Cette protéine présente une forte homologie de séquence avec la région centrale d'eIF4G et interagit avec les RRM 1 et 2 et le domaine C-terminal (PABC) de la PABP ([Roy et al., 2002](#)). A l'instar d'eIF4G, Paip1 interagit avec eIF4A mais ne peut lier eIF4E et une interaction potentielle avec eIF3 reste à démontrer ([Craig et al., 1998](#)). La protéine Paip2 est quant à elle un inhibiteur de l'activité de la PABP par sa liaison

avec les motifs RRM 2 et 3 et le domaine PABC. Cette protéine empêche ainsi la liaison entre Paip1 et la PABP, et déstabilise le complexe PABP:poly(A) ([Khaleghpour et al., 2001](#)) (**Figure 1.36 (A)**). Une étude récente suggère un mode d'action complémentaire impliquant une compétition avec eIF4G au regard de la liaison avec la PABP ([Karim et al., 2006](#)). Une séquence consensus suffisante pour la liaison par la PABP a pu être identifiée par alignement des séquences de Paip1 et Paip2 ([Kozlov et al., 2001](#)). La structure du domaine C-terminal de la PABP impliqué dans cette liaison a été résolue par RMN ([Kozlov et al., 2001](#)). Le domaine s'organise comme une pointe de flèche formée par 5 hélices α . Le rôle de Paip1 et Paip2 au sein du démarrage de la traduction reste cependant flou.

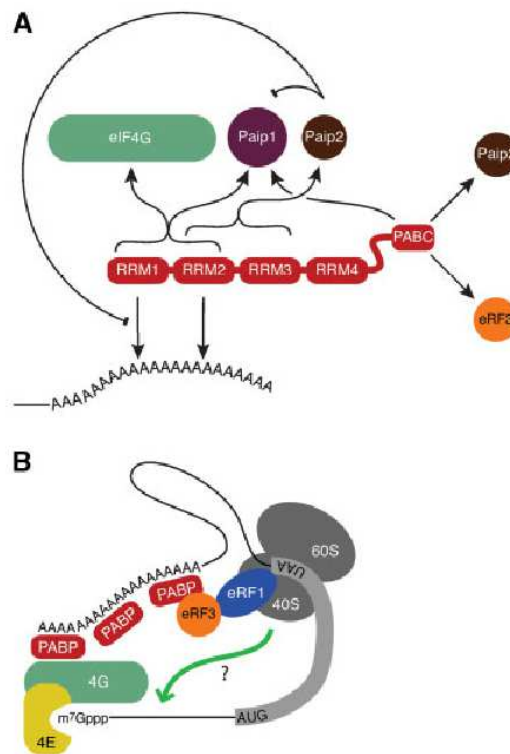


Figure 1.36 : Rôle de la PABP et la queue poly(A) dans la traduction d'après ([Preiss and M., 2003](#)). (A) Les différents domaines de la PABP sont indiqués. Les flèches représentent les liaisons observées. Les traits se terminant par un barre repèrent les liaisons empêchées par la liaison de Paip2. (B) Le modèle de recyclage d'un ribosome en posttermination.

Une autre interaction impliquant la PABP soulève de nombreuses questions. Le partenaire concerné, eRF3, est en effet impliqué dans la terminaison de la traduction ([Hoshino et al., 1999](#)). Le domaine C-terminal de cette protéine interagit avec eRF1, le facteur qui reconnaît les codons stop et permet l'éjection de la chaîne polypeptidique du ribosome ([Frolova et al., 1994](#)). eRF3 est nécessaire à l'activation de l'activité GTPasique d'eRF1 requise pour son activité ([Zhouravleva et al., 1995](#)). En revanche, le domaine N-terminal n'est pas requis pour l'étape de terminaison. Ce domaine interagit avec la PABP. La surexpression de ce domaine isolé inhibe la traduction *in vitro* par une voie indépendante de Paip1. De plus, cette inhibition interviendrait au cours d'une étape de recyclage du ribosome et pas sur des complexes nouvellement formés. L'interaction pourrait ainsi créer un lien entre un codon stop de l'ARNm et la queue poly(A). La région 3' non codante en aval du codon stop formerait alors une boucle et le ribosome pourrait passer à l'extrémité 5' par l'intermédiaire d'eIF4F et alors effectuer un nouveau cycle sur cet ARNm (**Figure 1.36 (B)**). Peu d'éléments permettent à ce jour de confirmer ce processus.

4 Démarrage par un Site Interne d'Entrée du Ribosome (IRES)

Le mode de démarrage de ces ARNm contenant un IRES est un mécanisme totalement indépendant de la présence d'une coiffe sur l'ARNm. La traduction effective des IRES repose sur sa structure secondaire et tertiaire, sur la participation de facteurs protéiques et sur la séquence complémentaire de l'ARNr 18S ([Chappell et al., 2000](#); [Le and Maizel, 1997](#); [Pilipenko et al., 2000](#)). Les facteurs de démarrage canoniques excepté le facteur eIF4E sont requis pour une traduction efficace. Néanmoins certains IRES viraux peuvent être traduits avec un nombre réduit de facteurs d'initiation. Notamment, l'ARN du virus de la paralysie du cricket peut s'associer à la sous-unité 40S et stimuler la formation d'un ribosome 80S sur un codon alanine permettant le démarrage de sa traduction par l'ARNt élongateur alanine sans intervention d'autres facteurs ([Pestova and Hellen, 2003](#)) (**Figure 1.37 (B)**).

Ce cas extrême est sensiblement différent des autres IRES viraux qui requièrent l'assistance de facteurs de démarrage et notamment d'eIF2 afin d'apporter l'ARNt initiateur au ribosome d'eIF3 pour stabiliser le complexe ternaire sur le ribosome et l'associer au 40S:mRNA ([Pestova et al., 1996a](#); [Sizova et al., 1998](#)). La présence des autres facteurs de démarrage (excepté eIF4E) stimule néanmoins la formation d'un ribosome 80S ([Pestova et al., 1996a](#)). La structure obtenue par cryomicroscopie électronique de l'IRES du virus de l'hépatite C lié à la sous-unité 40S révèle comment un IRES s'accapare la machinerie traductionnelle ([Siridechadilok et al., 2005](#)) (**Figure 1.37 (A)**). L'IRES apparaît sous une forme étendue. Il se fixe sur un site accessible au solvant, à un endroit cohérent avec la localisation du canal de liaison de l'ARNm au niveau du cou de la sous-unité 40S. Le domaine de l'IRES dépassant de la plateforme permet l'arrimage du facteur eIF3 pour stabiliser la liaison de l'IRES sur la sous-unité 40S (**Figure 1.37 (A)**). De plus, un autre domaine important pour l'activité de l'IRES modifie la conformation du canal de liaison de l'ARNm par un contact avec la tête, limitant ainsi son mouvement. La boucle terminale du segment contacte l'ARN viral simple brin codant et le maintient ainsi dans le site de décodage. La résolution d'une structure avec l'IRES du virus de la paralysie du cricket montre bien la différence entre ces systèmes : la configuration de la petite sous-unité ribosomale est modifiée de la même façon. En revanche, la localisation du facteur ainsi que certains changements de conformations typiquement rencontrés au sein des ribosomes 80S révèlent l'habileté de ce virus à utiliser la machinerie traductionnelle de son hôte sans aucune autre intervention. L'IRES ressemble alors étrangement à un facteur de démarrage composé d'ARN capable d'interagir avec les deux sous-unités ribosomales ([Spahn et al., 2004b](#)).

Une catégorie d'IRES reste très controversée. Il s'agit de ceux portés par des ARNm cellulaires qui seraient traduits dans des conditions où la synthèse protéique coiffe-dépendante est inhibée. Cette classe d'ARNm coderait pour des protéines impliquées notamment dans le contrôle du cycle cellulaire et l'apoptose. La difficulté à déterminer si ces ARNm sont effectivement traduits à l'aide d'IRES dans les conditions physiologiques contribue à la mise en doute de leurs existences ([Kozak, 2005](#)). Néanmoins, une étude récente chez la levure montre la nécessité pour la cellule d'exprimer des ARNm contenant de tels IRES en réponse à un stress afin d'assurer une adaptation aux stimuli ([Gilbert et al., 2007b](#)). Ce mécanisme fournirait un mode alternatif de démarrage et notamment dans le cas où eIF4E est piégé par ses partenaires et donc quand la traduction coiffe-dépendante est sévèrement limitée

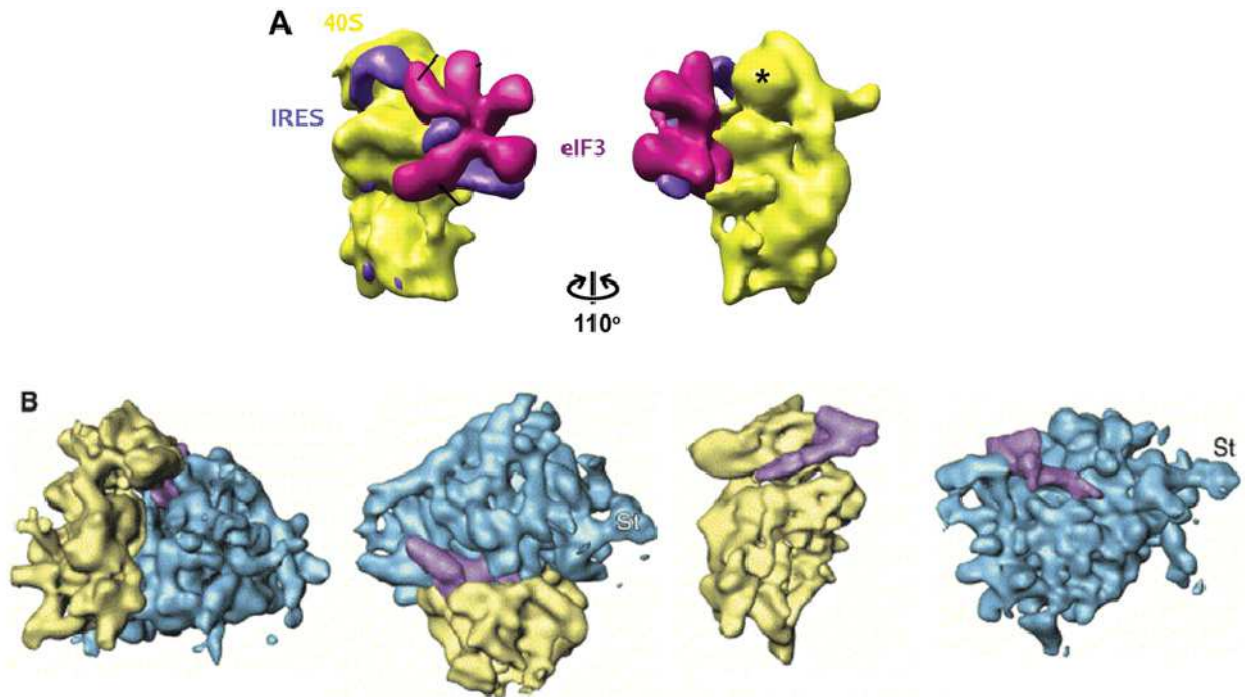


Figure 1.37 : Mode d'action des IRES du HCV et du CrPV. (A) Structure de la sous-unité 40S liée à l'IRES du HCV et à eIF3 d'après ([Siridechadilok et al., 2005](#)). (B) Structure du ribosome 80S lié à l'IRES du CrPV d'après ([Spahn et al., 2004b](#)). De gauche à droite, vue de côté, vue de dessus, vue sans la sous-unité 60S (bleue) et vue sans la sous-unité 40S (jaune). L'IRES est coloré en violet et la position du groupe de protéines P est indiqué (St).

PARTIE D LE FACTEUR EIF3

Identifié depuis les années 1970 en tant que stimulateur de la synthèse protéique, l'organisation et la fonction du facteur de démarrage eIF3, spécifique au système eucaryotique, commence à être comprise dans des systèmes simples. En effet, le rôle central qu'occupe ce facteur ainsi que les multiples réseaux d'interactions qu'il organise ont dans un premier temps limité la caractérisation de ce facteur multimérique d'environ 700kDa chez les mammifères.

I Composition d'eIF3 chez les Eucaryotes

Le facteur de mammifère apparaît comme le plus complexe des différents eIF3 eucaryotiques. Il se compose de 13 sous-unités différentes que la nomenclature unifiée définie en 2001 désigne de eIF3a à eIF3m, par ordre décroissant de taille ([Asano et al., 1997](#); [Browning et al., 2001](#); [Damoc et al., 2007](#); [Mayeur et al., 2003](#)) (**Tableau 1.2**). Précédemment, les sous-unités de mammifères étaient désignées en fonction de leurs masses ; par exemple eIF3a était connue comme p170. La caractérisation des sous-unités du facteur chez la levure avait mené à les nommer en tant que produit de gènes lorsque ceux-ci étaient définis; par exemple, eIF3a était appelée TIF32p ou en fonction de leurs masses. Une analyse comparative des séquences des différentes sous-unités a fait apparaître une conservation au sein des Eucaryotes et la nomenclature simplifie la description des homologies. Ainsi, le facteur de la levure *Saccharomyces cerevisiae*, ou levure de boulanger, apparaît beaucoup plus simple que le facteur humain. Il se compose de cinq sous-unités stœchiométriques, orthologues des cinq sous-unités de mammifère correspondantes : eIF3a, eIF3b, eIF3c, eIF3g et eIF3i. Les observations sur le facteur de levure capable d'assurer les mêmes fonctions que le facteur humain ont tout d'abord conduit à identifier un complexe eIF3 de « cœur » composé des sous-unités présentes chez la levure ([Phan et al., 1998](#)). En effet, ces sous-unités sont essentielles pour la traduction *in vivo* ([Greenberg et al., 1998](#); [Hanachi et al., 1999](#); [Naranda et al., 1997](#); [Valasek et al., 1998](#); [Verlhac et al., 1997](#); [Vornlocher et al., 1999](#)). Une sixième sous-unité orthologue : eIF3j, est une protéine non-essentielle, sous-stœchiométrique dont le rôle serait plutôt régulateur ([Nielsen et al., 2006](#); [Valasek et al., 1999](#)). Enfin, une protéine non homologue à la sous-unité humaine eIF3e et non essentielle nommée Pci8p présente un domaine conservé avec eIF3e qui lui permet de se lier au complexe eIF3 de la levure de boulanger au même titre que la sous-unité humaine *in vitro* et *in vivo*. Pci8p serait ainsi un orthologue divergent d'eIF3e ([Shalev et al., 2001](#)).

Chez la levure *Schizosaccharomyces pombe*, ou levure de fission, le facteur eIF3 se compose d'orthologues de la plupart des sous-unités humaines. Les sous-unités eIF3g et eIF3i sont néanmoins essentielles alors que les sous-unités eIF3d et eIF3e ne faisant pas partie du « cœur conservé » ne le sont pas ([Akiyoshi et al., 2001](#); [Bandyopadhyay et al., 2000](#); [Chen et al., 1999](#); [Crane et al., 2000](#)). En revanche, les sous-unités eIF3f et eIF3m ne possédant pas d'orthologues chez *S. cerevisiae* sont essentielles pour la traduction chez *S. pombe*. Ainsi, certaines sous-unités qui ne font pas partie du « cœur conservé » peuvent être requises pour assurer des activités d'eIF3 vitales ou des activités de régulation essentielles à la survie. De façon intéressante, il a pu être identifié chez ces organismes deux formes du complexe eIF3, chacune comprenant les cinq sous-unités de cœur, eIF3f et soit eIF3h et eIF3m, soit eIF3d et eIF3e ([Zhou et al., 2005](#)).

Nom de la sous-unité	Motif consensus	Nom du gène ^a	Humain	Blé	<i>A. thaliana</i>	<i>S. cerevisiae</i>
			MW ^b	MW ^b	MW ^c	MW ^c
eIF3a	PCI	<i>TIF32/RPG1</i> ^d	170	116	114	110
eIF3b	RRM	<i>PRT1</i>	116	83	82	90
eIF3c	PCI	<i>NIP1</i>	110	107	105	93
eIF3d	Aucun		66	87	66	Aucune
eIF3e	PCI	(<i>INT6</i>)	48	45	51	Aucune
eIF3f	MPN		47	34	32	Aucune
eIF3g	RRM, Doigt Zn	<i>TIF35</i>	44	36	33	33
eIF3h	MPN		40	41b	38	Aucune
eIF3i	répétitions WD	<i>TIF34(TRIP1)</i>	36	41a	36	39
eIF3j	Aucun	<i>HCR1</i> ^e	35	Aucune	Aucune	30
eIF3k	Aucun		28	28	25	Aucune
eIF3l	Aucun		Aucune	56	60	Aucune
eIF3m	PCI		42	Aucune	Aucune	Aucune

Tableau 1.2 : Nomenclature unifiée des sous-unités du facteur eIF3 (eukaryotic Initiation Factor 3) et leurs poids moléculaires (MW). ^a Les cases vacantes représentent les gènes ne comportant pas encore de nom. Les noms entre parenthèses font références aux gènes humains ; les autres proviennent de la levure. ^b Les poids moléculaires (kDa) ont été déterminés par migration sur SDS-PAGE. ^c Les poids moléculaires (kDa) sont calculés à partir des séquences prédites des protéines. ^d *RPG1* : Required for Passage through G1 phase ; ^e *HCR1* : High Copy Suppressor of Rpg1. Définition des motifs : PCI : Proteasome, COP9, Initiation factor 3 ; MPN : Mpr1, Pad1, N-terminal ; RRM : RNA Recognition Motif.

Une grande conservation dans les sous-unités du facteur de démarrage eIF3 est également observée au sein des génomes des plantes, de celui du blé à celui d'*Arabidopsis thaliana*. Ces facteurs sont très proches du facteur humain à la différence qu'il n'a pu être identifié d'orthologues des sous-unités eIF3j et eIF3m.

L'existence du complexe eIF3 de « cœur » justifie le choix de la levure *Saccharomyces cerevisiae* comme organisme modèle.

II Organisation du facteur eIF3 de *S. cerevisiae*

Le fruit du travail de l'équipe d'Alan Hinnebusch avec notamment Leo Valášek nous révèle l'architecture du complexe de la levure. Celle-ci a été déduite de nombreuses études de délétions de régions des sous-unités puis de l'analyse des sous-complexes formés *in vitro* et *in vivo* ([Asano et al., 1998](#); [Valasek et al., 2002](#); [Valasek et al., 2001](#); [Verlhac et al., 1997](#)).

Dans un premier temps, des résultats obtenus par des expériences de double-hybride chez la levure ont fourni les premières preuves génétiques d'une interaction entre eIF3i et eIF3b ([Verlhac et al., 1997](#)).

Ce type d'expérience requiert la construction de deux protéines hybrides. L'une est formée d'un domaine de liaison à l'ADN fusionné à une des protéines d'intérêt et l'autre est composée d'un domaine d'activation de la transcription fusionné à une autre protéine étudiée (**Figure 1.38**). Ces protéines hybrides sont exprimées dans une souche de levure comportant un ou plusieurs gènes rapporteurs régulés par le facteur de transcription activé en présence du domaine de liaison à l'ADN et du domaine d'activation de la transcription. Ainsi, si les protéines hybrides interagissent, elles formeront un activateur de la transcription fonctionnel en raison de la proximité du domaine de liaison de l'ADN et du domaine d'activation de la transcription. Les gènes rapporteurs seront alors effectivement exprimés. Usuellement, les gènes rapporteurs sont *lacZ* et/ou *HIS3* sous le contrôle de sites de liaison de la protéine Gal4. L'interaction ainsi identifiée entre eIF3i et eIF3b était également observée avec des fragments C-terminaux d'eIF3b.

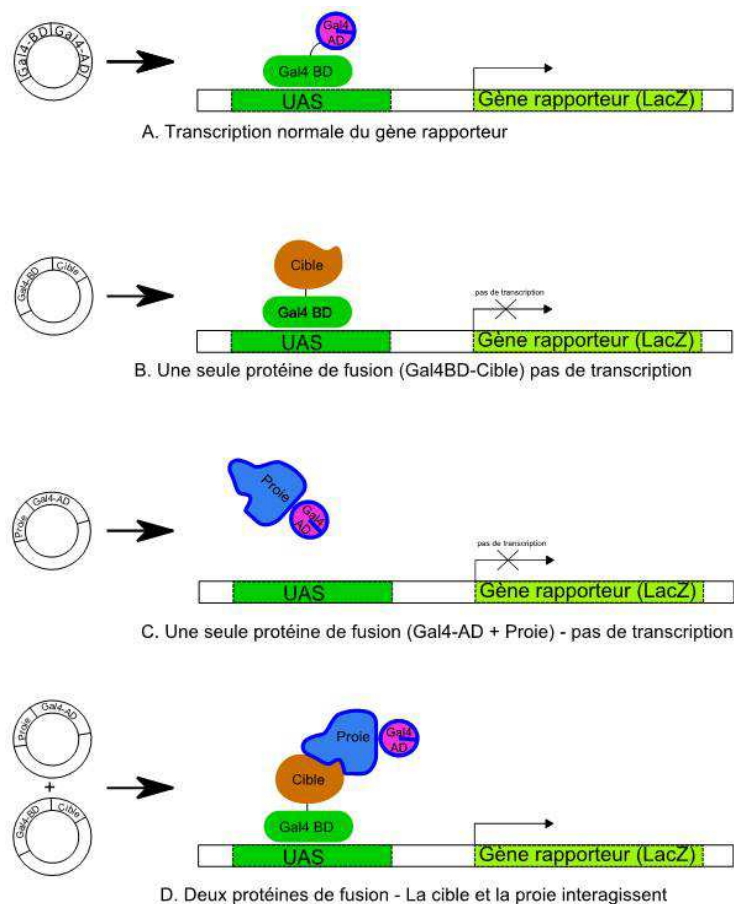


Figure 1.38 : Principe du double hybride avec le système Gal4 UAS-LacZ d'après (Anna K., http://en.wikipedia.org/wiki/Two_hybrid). UAS (Upstream Activating Sequence) est le domaine d'ADN où se fixe le domaine de liaison de Gal4 (Gal4-BD : Gal4-Binding Domain). Le domaine activateur de la transcription est noté Gal4-AD (Gal4-Activation Domain). La présence de deux protéines interagissant est nécessaire pour activer la transcription du gène rapporteur.

Cette étude a de plus soutenu l'importance d'eIF3i par des analyses de tétrades et des mutations sensibles à la température.

L'analyse de tétrade consiste à faire sporuler les levures et donc obtenir à partir d'une cellule diploïde quatre cellules haploïdes. Ainsi, si un allèle *TIF34* codant pour eIF3i est inactivé sur un chromosome, la cellule diploïde est toujours viable. Si l'inactivation sur des deux allèles est létale, deux des quatre cellules haploïdes ne pourront pas se développer. L'essentialité du gène codant pour eIF3i a ainsi pu être confirmée ([Naranda et al., 1997](#); [Verlhac et al., 1997](#)). De plus, la complémentation de la souche par un plasmide permettant d'exprimer eIF3i restaure la viabilité. Les auteurs ont également obtenu des mutations dans le gène *TIF34* rendant les levures sensibles à la température. En effet, ces souches croissent normalement à 24°C mais pas à 37°C. Le séquençage du gène d'eIF3i dans ces souches révèle trois mutations dans des motifs de type WD-40, identifiés dans la séquence d'eIF3i et conservés de la levure à l'humain ([Chen et al., 1995](#); [Verlhac et al., 1997](#)). Un motif WD-40 correspond à une région de la protéine qui débute par une séquence WD et qui contient 40 résidus. Nous reviendrons sur ces motifs dans le chapitre consacré à eIF3i. Une autre souche isolée contient deux mutations en dehors de ces régions. Néanmoins, il est conclu que les répétitions WD-40 doivent avoir une importance fonctionnelle. L'analyse des conséquences de ces mutations montre que le passage de la température permissive à la température non-permissive s'accompagne d'un arrêt de la synthèse protéique. Une analyse du profil des polysomes de ces souches mutantes a été entreprise. Une heure après le passage à la température non-permissive, le ratio de polysomes comparés aux monosomes est significativement réduit, signe d'un défaut dans le démarrage de la traduction ([Verlhac et al., 1997](#)).

Une analyse par le système de double hybride a ensuite été entreprise sur les cinq sous-unités de levure ([Asano et al., 1998](#)). Dans un premier temps, cela a mis en évidence des interactions entre eIF3b, eIF3g et eIF3i (**Figure 1.39**). En utilisant des versions raccourcies des différents facteurs, le site de fixation d'eIF3g et d'eIF3i a pu être limité aux 80 résidus C-terminaux d'eIF3b. De plus, le domaine de liaison d'eIF3i sur eIF3g serait compris dans la moitié N-terminale d'eIF3g, en dehors du domaine RRM C-terminal identifié. Ces interactions ont été confirmées *in vitro* par la technique de péchage par affinité (« pull-down »). Cette méthode consiste à fusionner une protéine à une étiquette pour faciliter sa purification. Dans ce cas, il s'agit de la GST (Glutathione S-Transferase). La protéine est ensuite immobilisée sur une résine à forte affinité pour la GST (Glutathione Sepharose). Des extraits bruts ou des protéines purifiées éventuellement marquées sont mis en contact avec la résine et le complexe associé après lavage de la résine est ensuite analysé pour identifier les protéines retenues. Par cette méthode, il apparaît que les 42 résidus C-terminaux d'eIF3b suffisent à la liaison d'eIF3i et de la moitié N-terminale d'eIF3g. L'extension de cette région aux 80 résidus C-terminaux assure néanmoins la liaison optimale d'eIF3i et eIF3g.

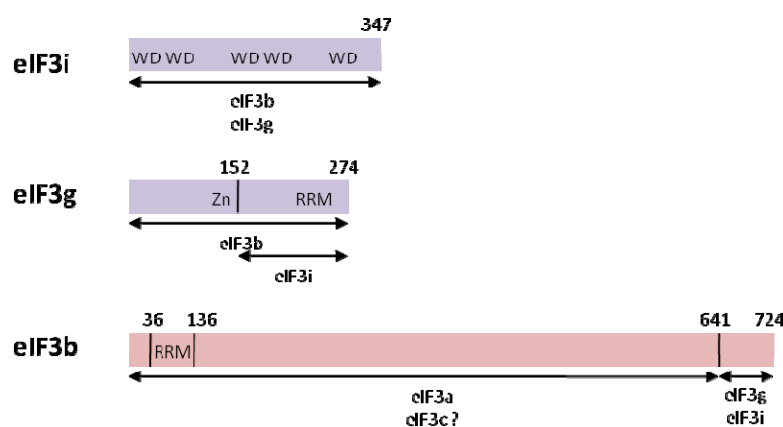


Figure 1.39 : Représentation schématique des résultats obtenus par ([Asano et al., 1998](#)). Les répétitions WD, les domaines de liaison à l'ARN (RRM) et le domaine de liaison du zinc (Zn) sont indiqués. Les doubles flèches indiquent les domaines d'interaction.

Les protéines eIF3a et eIF3c ont été testées dans la technique de GST-« pull-down » afin d'identifier des protéines partenaires n'apparaissant pas par le double hybride. Des formes tronquées de la région C-terminale d'eIF3b peuvent ainsi interagir avec eIF3a montrant que le site d'interaction serait en N-terminal de celui d'eIF3g et eIF3i. De plus, eIF3a contacte eIF3c dans ce type d'expérience même si cette interaction paraît plus faible. Cette interaction pourrait cependant être gênée par la protéine de fusion.

La suite des investigations a porté sur la caractérisation des interactions entre eIF3a, eIF3b, eIF3c et eIF3j ([Valasek et al., 2001](#)) (**Figure 1.40**). En premier lieu, il convient de remarquer qu'eIF3j (codé par le gène HCR1) est fortement homologue (25%) à un domaine interne d'eIF3a (résidus 490-790), nommé HLD (HCR1-Like Domain). Par GST-« pull-down », le site de liaison minimal d'eIF3b requis pour la liaison d'eIF3a et d'eIF3j a pu être déterminé. Il réside dans les 136 premiers résidus d'eIF3b et comprend un motif RRM. La liaison d'eIF3j requiert moins de résidus et nécessite uniquement le motif RRM. Une interaction simultanée d'eIF3a et eIF3j avec le motif RRM d'eIF3b a pu être mise en évidence. Le site minimal de liaison d'eIF3b sur eIF3a a de même été identifié. La fixation d'eIF3b nécessite le domaine HLD et le domaine C-terminal (résidus 790-964). En outre, la moitié N-terminale d'eIF3a permet de lier les facteurs eIF3c et eIF3j. Néanmoins, une troncature C-terminale d'eIF3a n'interagit pas aussi fortement avec eIF3j que la protéine entière. Cela indique une région C-terminale assurant des contacts supplémentaires ou bien que cette région assure une conformation optimale de la région N-terminale de la protéine favorisant ainsi l'interaction. L'importance du motif RRM d'eIF3b en tant que zone d'interaction avec eIF3a est soutenue par l'observation qu'une délétion de ce domaine induit un phénotype de croissance lente *in vivo* ([Evans et al., 1995](#)). Ainsi, la protéine mutante pourrait lier eIF3g et eIF3i mais pas eIF3a et alors les séquestrer, empêchant la formation d'un complexe intact. Le phénotype a pu être partiellement levé par la surexpression d'eIF3i mais pas d'eIF3g, suggérant qu'eIF3i serait le facteur limitant dans ces souches. En fusionnant l'allèle du gène d'eIF3b tronqué de son domaine N-terminal avec une étiquette, un complexe avec eIF3g et eIF3i est isolé *in vivo*. En revanche, il n'a pu être observé de fixation d'eIF3a, eIF3c ou eIF3j à la différence de ce qui est observé avec l'allèle sauvage du gène d'eIF3b.

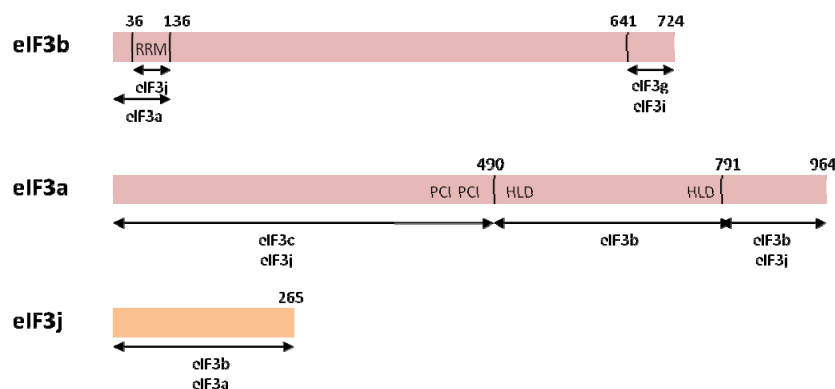


Figure 1.40 : Représentation schématique des résultats obtenus par ([Valasek et al., 2001](#)). Les domaines PCI et HCR1-like (HLD) sont indiqués.

Valášek et al. ont ensuite cherché à caractériser plus finement la surface d'interaction entre eIF3a et eIF3c par GST-« pull-down » *in vitro* ([Valasek et al., 2002](#)). Le site minimal d'eIF3a requis pour sa liaison à eIF3c se situe entre les résidus 200 et 600. Sur eIF3c, la région comprise entre les résidus 157 et 370 permet la liaison d'eIF3a. Un nouveau contact entre eIF3b et eIF3c a finalement été montré. En effet, une troncature d'eIF3c après le résidu 371 ne permet de former un complexe pentamérique dont les sous-unités eIF3b, eIF3g et eIF3i sont largement sous-stœchiométriques. En revanche, une troncature plus modérée après le résidu 570 est nécessaire pour former un complexe eIF3 intègre *in vivo*. Ainsi, il a été montré que la partie 371-570 d'eIF3c contient un site additionnel d'interaction entre eIF3b et eIF3c. Cette région d'eIF3c est en contact avec la région 551-641 d'eIF3b.

L'ensemble du réseau d'interaction mis en évidence par ces études est illustré dans la **Figure 1.41**.

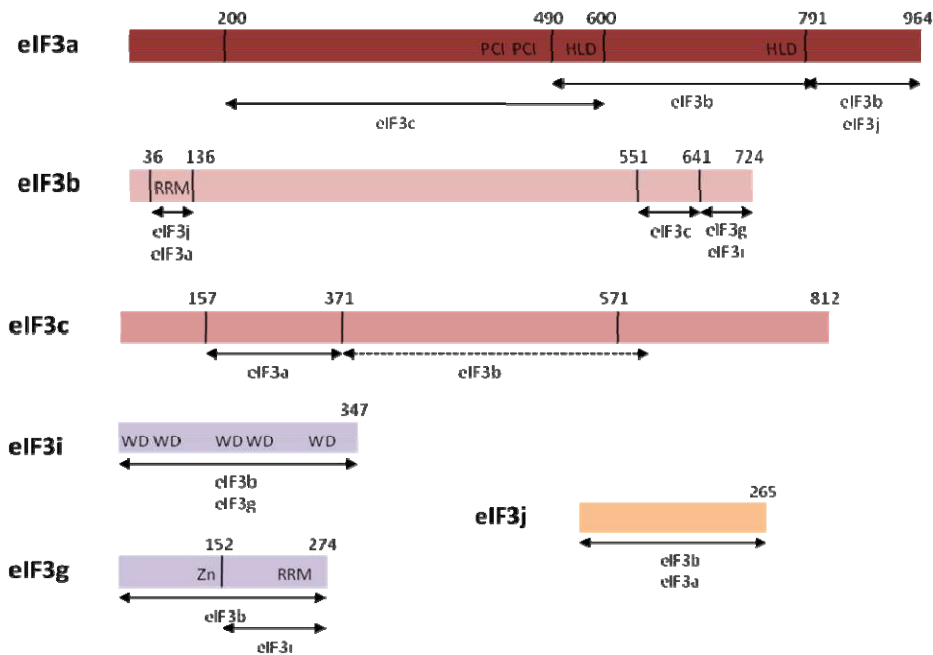


Figure 1.41 : Représentation schématique de l'ensemble des résultats obtenus sur les interactions entre les sous-unités du facteur eIF3.

III De l'organisation à la fonction

1 Le complexe Multi-Facteur (MFC) de *S. cerevisiae*

La caractérisation du complexe eIF3 a notamment été corrélée à la découverte du complexe multifacteur (MFC) dont il est l'élément architectural principal (**Figure 1.42**). En effet, le complexe eIF3 co-purifie avec les facteurs de démarrage eIF1 et eIF5 ([Naranda et al., 1997](#); [Phan et al., 1998](#)). L'interaction entre eIF3c et eIF1 a également été observée dans un système de double hybride ([Asano et al., 1998](#)). La partie C-terminale d'eIF3c (1-156) permet de former un complexe avec la partie C-terminale d'eIF5 contenant les boîtes AA impliquées dans sa liaison à eIF2 β mais également le facteur eIF1. Les fixations conjointes ont été observées par double hybride et GST-« pull-down » ([Asano et al., 2000](#); [Asano et al., 1999](#); [Phan et al., 1998](#)). Conformément à ce schéma, un sous-complexe formé d'eIF3a, eIF3b et eIF3c est capable de lier *in vivo* eIF1 et eIF5 ([Phan et al., 2001](#)). De plus, le domaine C-terminal du facteur de démarrage eIF5 peut se lier simultanément à eIF2 β , eIF3c et eIF1 suggérant que le facteur puisse lier ces partenaires au sein du MFC ([Asano et al., 2000](#); [Asano et al., 1999](#); [Singh et al., 2004](#)). Ce complexe MFC a en effet pu être purifié à partir d'un extrait brut sans être associé au ribosome. Il contient un complexe ternaire intact ([Asano et al., 2000](#)). De plus, des mutations dans les boîtes AA d'eIF5 altèrent son réseau d'interaction avec eIF2 β et eIF3c *in vitro* et réduisent l'association entre eIF2, eIF3 et eIF5 dans les extraits bruts. Cette mutation provoque une sensibilité à la température qui résulte d'une diminution du démarrage de la traduction. En effet, une réduction des pics de polysomes par rapport aux pics de monosomes est observée alors que les monosomes analysés correspondent à des ribosomes 80S sans complexe ternaire. En outre, un phénotype sauvage peut être restauré par la surexpression du complexe ternaire indiquant que l'interaction entre eIF5 et eIF2 est importante *in vivo* ([Asano et al., 2000](#); [Asano et al., 1999](#)). Finalement, eIF1 peut interagir avec eIF2 β en plus d'eIF5 et eIF3c densifiant ainsi le réseau d'interactions ([Singh et al., 2004](#)).

La prédiction d'une interaction entre eIF2 et eIF3 pontée par eIF1 et eIF5 a pu être confirmée par « pull-down » *in vivo* en purifiant les complexes formés par des versions tronquées d'eIF3c fusionnées à une étiquette octahistidine. Ainsi, une forme d'eIF3c dénuée de sa partie N-terminale permet de purifier un complexe eIF3 avec ses cinq sous-unités mais les facteurs eIF1, eIF2 et eIF5 en sont absents alors que cette partie N-terminale seule co-purifie avec ces facteurs sans les autres composants du complexe eIF3 (**Figure 1.42**). Cette étude a de plus fait apparaître l'existence dans le domaine C-terminal d'eIF3a de sites de liaison d'eIF1 et eIF2β correspondant aux domaines 791-964 et 642-791 respectivement. La partie C-terminale (791-964) peut ainsi lier eIF2β *in vitro* et eIF2 *in vivo* et former un sous complexe sans les autres partenaires du MFC. Alors que la délétion de ce domaine est létale, surexprimer eIF3a tronquée de ce domaine l'empêche de lier eIF2 et confère un phénotype de croissance lente. A l'instar de la mutation d'eIF5, le phénotype sauvage est restauré par la surexpression du complexe ternaire justifiant l'importance de l'interaction entre eIF3 et eIF2 ([Valasek et al., 2002](#)).

Dans ce contexte, un fragment interne d'eIF3a, sans site de liaison à eIF2 mais laissant intact le site de liaison d'eIF3c, résulte dans la formation d'un complexe MFC intact grâce à l'interaction entre eIF3c et eIF2. En revanche, la forme tronquée d'eIF3a uniquement dénuée du site de liaison d'eIF2 empêche la liaison d'eIF2 à eIF3. Il est proposé que le domaine N-terminal d'eIF3a déstabilise l'association d'eIF2 au sein du MFC. Ainsi, l'interaction entre eIF2 et le domaine C-terminal d'eIF3a est nécessaire afin de maintenir eIF2 au sein du MFC ([Valasek et al., 2002](#)).

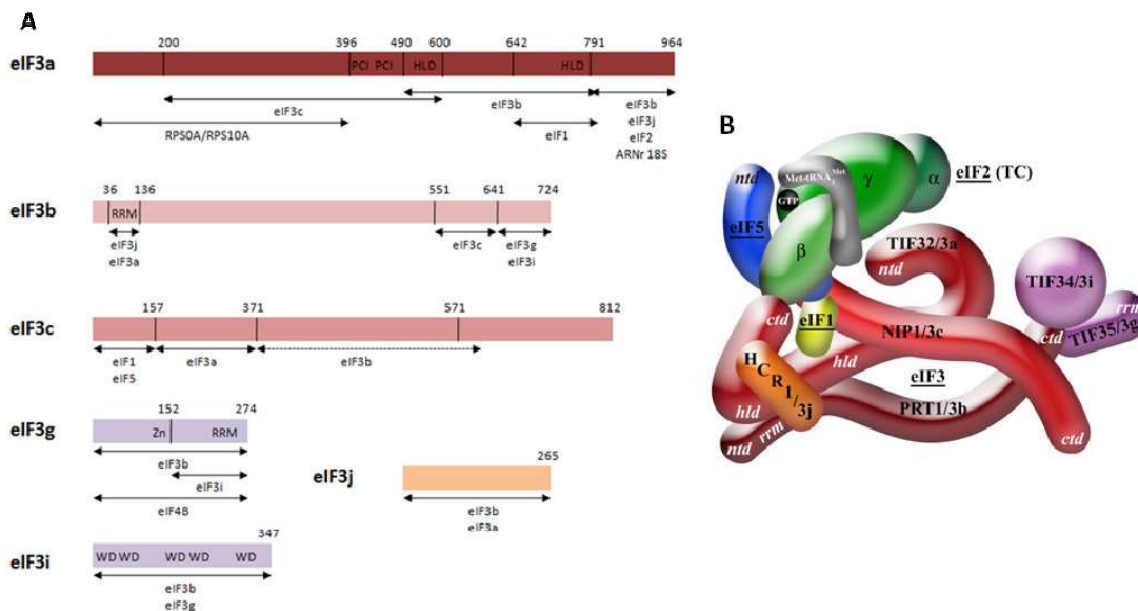


Figure 1.42 : Organisation du MFC par le facteur de démarrage eIF3 d'après ([Valasek et al., 2002](#)). (A) Représentation schématique des interactions répertoriées pour les différentes sous-unités du facteur eIF3. Les interactions d'eIF3a avec l'ARNr et les protéines ribosomales S0A et S10A sont indiquées. (B) Vue schématique du complexe multifacteurs.

De façon remarquable, la reconstitution de complexe de prédémarrage *in vitro* révèle qu'un sous-complexe formé des sous-unités eIF3a, eIF3b et eIF3c peut stimuler la liaison du complexe ternaire et d'un ARNm à la sous-unité ribosomale 40S avec la même efficacité que le complexe eIF3 sauvage ([Phan et al., 2001](#)). En revanche, un sous-complexe comprenant les sous-unités eIF3b, eIF3i et eIF3g ne possède pas une telle activité. Ceci incite à réduire le « cœur » fonctionnel d'eIF3 aux sous-unités eIF3a, eIF3b et eIF3c. Il faut cependant remarquer que les sous-unités eIF3i et eIF3g sont essentielles ([Hanachi et al., 1999](#); [Verlhac et al., 1997](#)). Ces sous-unités doivent donc intervenir à d'autres étapes de la traduction.

In vitro, la liaison d'eIF2 β avec eIF5 stimule la liaison d'eIF5 à eIF3c. L'interaction eIF2 β -eIF5 pourrait ainsi nucléer le sous-complexe eIF2-eIF3-eIF5 du MFC. Ainsi, la délétion des 10 résidus C-terminaux d'eIF5 altère sa liaison à eIF2 β *in vitro* mais également les interactions eIF3-eIF5 et eIF2-eIF3 *in vivo* ([Singh et al., 2004](#)). Le phénotype sensible à la température associé à cette mutation est de plus partiellement complété par la surexpression du complexe ternaire ([Singh et al., 2005](#)). Par un modèle de structure reposant sur l'homologie entre le domaine HEAT eIF2B ϵ et ce même domaine d'eIF5, il a été proposé que cette délétion altère uniquement le site de liaison d'eIF2 β sans modifier la liaison d'eIF3c ([Boesen et al., 2004](#); [Singh et al., 2004](#)). La mutation de résidus acides situés à l'extrémité C-terminale empêche également la liaison d'eIF2 β , précisant le domaine d'interaction. De même, la mutation d'un groupe de résidus basiques sur une face différente de la structure a permis de prédire cette zone comme le site de liaison d'eIF1 et eIF3c. En outre, une analyse de la séquence du domaine N-terminal d'eIF3c révèle la présence de régions acides ou basiques conservées. Ces régions pourraient être impliquées dans la liaison des motifs acides d'eIF5 et des régions acides et basiques d'eIF1 ([Asano et al., 2000](#); [Fletcher et al., 1999](#)).

Les effets sur la formation du complexe eIF2:eIF3:eIF5 *in vivo* sont plus complexes à interpréter et révèlent certainement des effets de couplages entre les différents partenaires de la machinerie ([Yamamoto et al., 2005](#)).

2 La formation du MFC stimule la formation du complexe 43S

La découverte du MFC chez la levure laisse penser qu'il serait recruté sous cette forme au ribosome par un mécanisme hautement coopératif (**Figure 1.43**). Ainsi, les perturbations d'interactions clés pour la formation du complexe devraient avoir des conséquences sur la formation du complexe 43S. Remarquablement, la mutation prt1-1 produit un phénotype de croissance lente et une réduction du démarrage de la traduction *in vivo* ([Hartwell and McLaughlin, 1969](#)). Le phénotype serait lié à un défaut de liaison de tous les composants du MFC sur la sous-unité 40S ([Phan et al., 2001](#)). De même, la mutation dans les boîtes AA d'eIF5 induit un défaut de recrutement du complexe ternaire au ribosome *in vivo* ([Asano et al., 2001b](#)). En revanche, la surexpression dans cette souche de la partie N-terminale d'eIF3a (dénuée de site d'interaction avec eIF2) réduit la liaison d'eIF2 et eIF3 au complexe 43S ([Nielsen et al., 2004](#)). Ainsi, il apparaît que les contacts entre eIF2 et eIF3 par l'intermédiaire d'eIF5 ou eIF3a forment des contributions additives *in vivo*. En effet, la formation d'un sous-complexe formé par la région N-terminale d'eIF3c et ses partenaires eIF1, eIF2 et eIF5 ne peut se lier efficacement à la sous-unité 40S ([Valasek et al., 2003](#)). Ce sous-complexe diminue la fixation d'eIF2 sur la sous-unité 40S *in vivo* et le phénotype est supprimé par la surexpression du complexe ternaire ([Nielsen et al., 2004](#); [Valasek et al., 2002](#)). Enfin, la mutation des résidus conservés de la partie N-terminale d'eIF3c induit un phénotype similaire à celui observé lorsque la partie N-terminale d'eIF3c est surexprimée. De plus, le phénotype est supprimé par la surexpression du complexe ternaire. Il résulte d'une liaison plus faible d'eIF3c à eIF1 et eIF5 *in vitro* et diminue la quantité d'eIF2 liée au ribosome *in vivo* ([Valasek et al., 2004](#)).

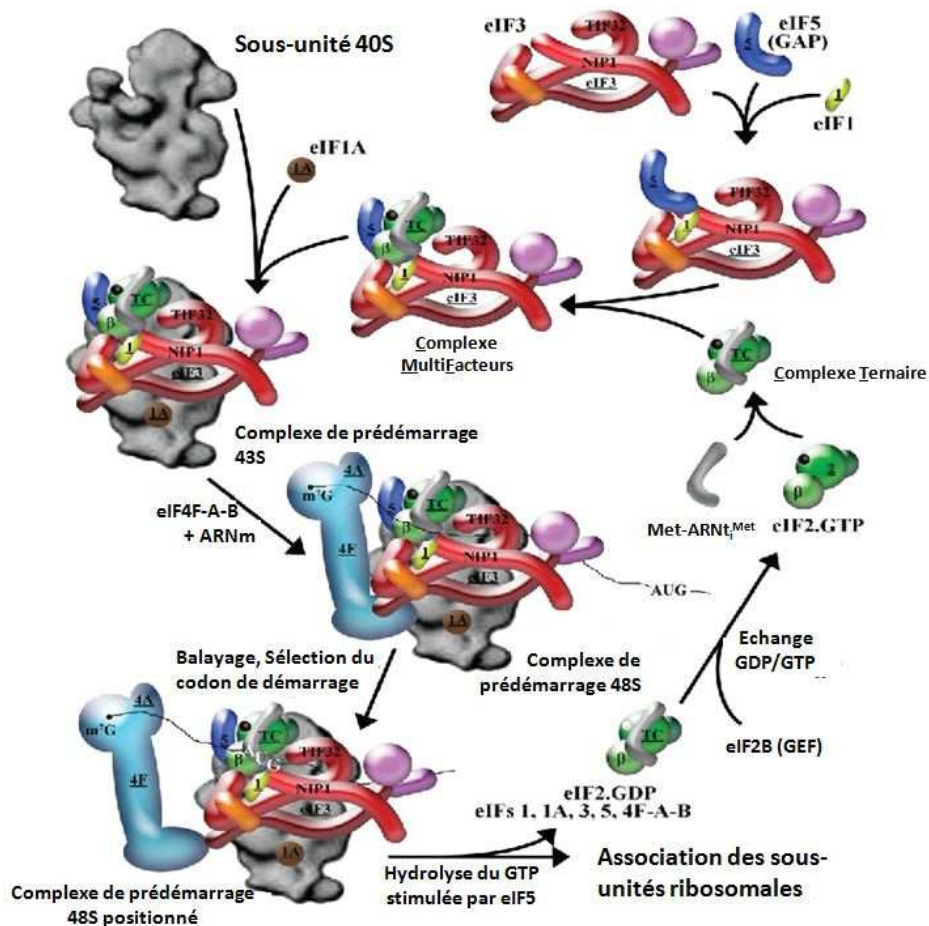


Figure 1.43 : Modèle proposé par Alan Hinnebusch et al. (<http://www.biomed.cas.cz/mbu/lrge/research.html>).

Une étude récente s'est proposé de montrer l'importance des différents composants du MFC sur la formation du complexe de prédémarrage 43S *in vivo*. Pour cela, les auteurs ont construit des souches comportant des « dégrons » sensibles à la température (notés td) sur des allèles codant pour des composants du MFC et notamment eIF2 β , eIF5, eIF4G et un double mutant eIF3a et eIF3b. Chaque allèle td code pour l'ubiquitine et une moitié dihydrofolate réductase thermolabile fusionnés en N-terminal du facteur de démarrage. Les allèles td sont sous contrôle d'un promoteur cuivre-dépendant et intégrés dans le chromosome de manière à annihiler l'allèle sauvage. Les souches mutantes td expriment également l'ubiquitine ligase sous contrôle d'un promoteur inductible au galactose. Ainsi, le passage des cellules d'un milieu à 25°C contenant du cuivre vers un milieu à 36°C sans cuivre contenant du galactose réprime la synthèse de ces facteurs et adresse les protéines ubiquitinées à une dégradation par le protéasome. Il a ainsi pu être constaté qu'aucune des mutations testées n'induit une perte complète du MFC, indiquant que ces facteurs ne sont pas entièrement interdépendants, mais toutes diminuent l'interaction du MFC avec la sous-unité 40S. L'étude montre l'importance *in vivo* de la présence des facteurs eIF2 et eIF3 dans le recrutement de l'ARNm. De plus, l'effet d'eIF5 dans la conversion du complexe 43S en complexe 48S est également confirmé. Les auteurs montrent que la délétion du facteur eIF4G n'empêche pas la formation d'un complexe 80S mais ces résultats se doivent d'être confirmés ([Jivotovskaya et al., 2006](#)). En effet, cette observation constituerait une forte différence entre les systèmes de levure et de mammifère. Chez ces derniers, le recrutement de l'ARNm au ribosome reposerait sur une interaction directe entre eIF4G et eIF3 ([Korneeva et al., 2000](#); [Lamphear et al., 1995](#)). Dans le cas de la levure, le modèle actuel propose que le recrutement d'eIF4G se produise via des interactions pontées par les facteurs eIF1 et eIF5 ([Asano et al., 2001b](#)).

IV Le rôle d'eIF3 dans la régulation de la traduction et le cancer

1 eIF3 et la régulation de la traduction

Nous avons vu quelques exemples de régulation de l'expression génétique au niveau du démarrage de la traduction par l'intermédiaire de facteurs de démarrage. Le rôle central qu'occupe le facteur eIF3 au sein du démarrage ainsi que quelques éléments expérimentaux suggèrent une intervention de ses sous-unités dans des processus de régulation de la traduction, et plus particulièrement, d'une sous-classe d'ARNm. En effet, il a été remarqué que le facteur eIF3 utilisé dans un système d'expression *in vitro* stimule différemment la traduction de deux ARNm différents ([Heywood et al., 1974](#)). De plus, la composition du facteur, notamment au regard de la présence de la sous-unité eIF3a du facteur humain ne semble pas modifier la capacité du complexe eIF3 à stimuler la formation d'un complexe de prédémarrage ([Chaudhuri et al., 1997](#)).

Les recherches menées sur le complexe eIF3 de *S. pombe* ont montré l'existence de deux sous-complexes. Tous deux comprennent les cinq sous-unités orthologues de la levure *S. cerevisiae* et eIF3f. L'un contient en plus les sous-unités eIF3h et eIF3m, appelé sous-complexe eIF3m. Le second, nommé sous-complexe eIF3e, contient quant à lui eIF3d et eIF3e. Le sous-complexe eIF3m s'associe avec l'ensemble des ARNm cellulaires alors que le sous-complexe eIF3e stimulerait la traduction d'un jeu d'ARNm beaucoup plus restreint. Ainsi, la composition du facteur eIF3 pourrait modifier sa fonctionnalité et donc agir sur le démarrage de la traduction de différentes espèces d'ARNm (Zhou et al., 2005).

A la différence de la sous-unité eIF3a de la levure de boulanger, essentielle pour la traduction *in vitro* et *in vivo*, l'altération de l'expression de la sous-unité de Mammifère par l'utilisation d'ADN complémentaire (ADNc) anti-sens dans des cellules cancéreuses provoque une diminution de l'ordre de 15-20% de la traduction globale ([Dong and Zhang, 2003](#); [Valasek et al., 1998](#)). Cette observation suggère que la sous-unité eIF3a de Mammifère n'est pas essentielle pour la biosynthèse protéique. Une altération de son expression est en fait corrélée avec des modifications d'expression de protéines impliquées dans le cycle cellulaire, indiquant un rôle possible de régulation ([Dong et al., 2004](#); [Dong and Zhang, 2003](#)). Le clivage d'eIF3a par le protéasome affecte l'assemblage de complexe de prédémarrage sur un nombre défini et non la population globale d'ARNm cellulaires et viraux dans un système de traduction reconstitué *in vitro* ([Baugh and Pilipenko, 2004](#)). On peut alors imaginer, à l'instar de la levure de fission, l'existence de deux sous-complexes contenant ou non eIF3a qui permettraient l'expression de différentes classes d'ARNm.

Environ 20% de la sous-unité eIF3a de mammifère sont associés aux membranes par fixation sur des filaments d'actine. Cette sous-population n'est pas phosphorylée alors que la protéine cytoplasmique est phosphorylée sur des résidus sérines et thréonines ([Pincheira et al., 2001a](#)). De plus, il a pu être observé une association des sous-unités eIF3a de Mammifère et de la levure de boulanger avec des microtubules ([Hasek et al., 2000](#); [Shanina et al., 2001](#)). Quatre signaux bipartites de localisation nucléaire ont pu être identifiés au sein de la sous-unité de Mammifère, impliquant que la sous-unité pourrait entrer dans le noyau après une protéolyse limitée ([Chudinova et al., 2004](#)). Ainsi, un rôle putatif de régulation de la biosynthèse protéique localisée peut-être attribué à eIF3a qui permettrait de contrôler le niveau d'expression de protéines spécifiques. Les investigations en cours sur cette protéine devraient nous permettre de mieux comprendre les différents niveaux d'action. Notamment, son implication dans les processus de phagocytose de cellules apoptotiques et de différenciation cellulaire ouvre de nombreuses perspectives ([Liu et al., 2007](#); [Nakai et al., 2005](#)).

La sous-unité eIF3h chez *A. thaliana* n'est pas strictement requise pour la traduction et son absence ne semble pas perturber la stabilité du complexe eIF3. De plus, son niveau d'expression semble corrélé au niveau de traduction de certains ARNm, suggérant un mode d'action similaire à celui d'eIF3a ([Kim et al., 2004](#)). Ces observations préliminaires appellent clairement des études supplémentaires.

L'apoptose s'accompagne d'une régulation forte de la traduction des ARNm. Alors que la traduction de la plupart des ARNm est inhibée, la traduction d'ARNm important pour l'exécution de la phase apoptotique est maintenue ou stimulée ([Clemens et al., 1998](#)). L'interaction d'eIF3f avec un médiateur de l'apoptose ainsi que la forme clivée de la sous-unité eIF3j, ne pouvant plus se lier à la sous-unité 40S, observées dans des cellules apoptotiques indiquent que ces deux protéines pourraient représenter des acteurs importants du processus ([Bushell et al., 2000](#); [Fraser et al., 2004](#); [Shi et al., 2003](#)).

L'infection virale est également couplée à un arrêt de la synthèse protéique des ARNm de l'hôte afin de traduire les ARN viraux plus efficacement. L'infection par le virus de l'herpès simplex (HSV) produit une protéine chargée de réguler la synthèse des protéines virales et celle de la cellule hôte. Cette protéine interagit avec eIF3b et eIF3f indiquant un possible mode d'action ([Fontaine-Rodriguez et al., 2004](#)). Plusieurs cas analogues ont été rapportés. L'infection de la souris par des virus ou d'autres stress cellulaires conduisent à l'expression des protéines Mup54 et Mup56 qui interagissent avec les extrémités N et C-terminales d'eIF3c, empêchant alors son interaction avec le complexe eIF4F ([Hui et al., 2005](#)). L'homologue humain de Mup56, Hup56, se lie en revanche à eIF3e induisant certainement l'altération d'une autre interaction cruciale ([Guo et al., 2000](#)). Enfin, eIF3d permettrait à la protéine VPg des calcivirus, liée covalamment à l'extrémité 5' de son ARN génomique, d'utiliser la machinerie hôte en se fixant à cette sous-unité du complexe eIF3 ([Daughenbaugh et al., 2003](#); [Herbert et al., 1997](#)).

La sous-unité eIF3g pourrait avoir un rôle dans le processus de redémarrage comme l'indiquent des études menées sur la traduction d'ORF d'ARNm polycistroniques chez les Eucaryotes. Ce mécanisme, décrit dans la partie C de l'introduction dans le cas de la traduction du gène *GCN4* chez la levure, est également utilisé par le pararetrovirus du virus de la mosaïque du chou-fleur (CaMV) ([Futterer and Hohn, 1991](#); [Hinnebusch, 1997](#)). L'ADN double brin de ce virus est transcrit par l'ARN-polymérase II en deux ARNm. Le premier ARNm dit 35S est redondant en sa partie 5' terminale et contient plusieurs ORF. Le second produit une protéine permettant d'activer l'expression efficace des différentes ORF majeurs de l'ARN 35S : TAV (Transactivator/viroplasmin) ([Bonneville et al., 1989](#); [Gowda et al., 1989](#)). La protéine TAV interagit avec eIF3g afin de se lier à eIF3 et de former un complexe TAV:eIF3:40S puis TAV:eIF3:80S ([Park et al., 2004](#); [Park et al., 2001](#)). Or ce site de liaison recouvre le site de fixation d'eIF4B sur eIF3g et pourrait indiquer comment le virus utilise la machinerie hôte afin de stimuler un redémarrage après la traduction de son premier ORF.

2 eIF3 et le cancer

Des études relativement récentes ont clairement établi un lien entre la régulation de la traduction par les facteurs du démarrage et l'oncogenèse. Les premières études indiquent que des expressions altérées de facteurs de démarrage pourraient engendrer des changements dans l'efficacité de la traduction d'ARNm spécifiques de manière à stimuler l'expression d'un jeu restreint d'ARNm. Ces derniers, traduits normalement à un niveau basal, coderaient pour des protéines-clés impliquées dans la croissance cellulaire, la survie cellulaire et le développement de cellules malignes ([De Benedetti and Graff, 2004](#); [Donze et al., 1995](#); [Fukuchi-Shimogori et al., 1997](#)). Plus récemment, des études ont montré des niveaux d'expression altérés de plusieurs sous-unités du facteur eIF3 dans des tumeurs malignes.

Notamment, le niveau d'expression d'eIF3a est augmenté dans plusieurs types de cellules et de tissus cancéreux. Premièrement identifié dans des tissus de cancer du sein, un niveau élevé d'eIF3a a été constaté dans des cancers du poumon, cervical, de l'œsophage et du colon ([Bachmann et al., 1997](#); [Chen and Burger, 1999, 2004](#); [Dellas et al., 1998](#); [Pincheira et al., 2001b](#)). De plus, les niveaux élevés sont plus fréquemment observés à des stades précoces plutôt qu'aux stades avancés des tumeurs ([Chen and Burger, 1999, 2004](#); [Dellas et al., 1998](#)). L'implication d'eIF3a dans l'oncogenèse se trouve renforcée par la réversibilité du phénotype malin de cellules du poumon cancéreuse par l'utilisation d'ADNc anti-sens, suggérant qu'eIF3a assurerait la maintenance du phénotype ([Dong et al., 2004](#)). Combinée à son implication dans la régulation de la traduction, la sous-unité eIF3a pourrait ainsi moduler la traduction d'une sous-classe d'ARNm qui code pour des protéines importantes pour la croissance cellulaire et l'oncogenèse.

De même, la sous-unité eIF3h voit son expression augmentée dans des cancers de la prostate, du sein et du foie ([Nupponen et al., 1999](#); [Okamoto et al., 2003](#); [Savinainen et al., 2004](#)). Il est proposé que la surexpression d'eIF3h altère la traduction d'ARNm codant pour des protéines importantes pour la croissance cellulaire.

L'expression de la sous-unité eIF3i est augmentée en réponse à un traitement au cadmium, un carcinogène humain ([Joseph et al., 2002](#)). De plus, la surexpression de cette protéine dans des cellules saines les transforme en cellules malignes alors que l'utilisation d'un ARN anti-sens annule à la fois la surexpression de la protéine et le phénotype induits par le cadmium ([Lei et al., 2002](#)). Or ce facteur humain, également identifié comme la protéine TRIP-1 (β -propeller), peut agir en tant que substrat des récepteurs TGF- β de type II. Ce récepteur fait partie d'une famille de facteurs de croissance : Transforming Growth Factors- β . L'activation de ce récepteur par la liaison de ligands induit une cascade de réactions, dites « voie TGF- β », aboutissant à une répression de l'activation de la transcription ([Choy and Derynck, 1998](#); [Matsuda et al., 2000](#)). Ainsi, la surexpression d'eIF3i réprime la stimulation de la transcription par la voie TGF- β et pourrait ainsi conduire à une croissance non contrôlée et à la formation de cancers.

Le gène codant pour la sous-unité eIF3e présente un site d'insertion fréquent pour le virus de tumeur mammaire ([Marchetti et al., 1995](#)). L'insertion du virus induit la production d'un ARNm codant pour une forme tronquée en C-terminal d'eIF3e. La surexpression dans des cellules saines de cette protéine engendre un phénotype de développement malin *in vitro* et *in vivo* alors que la surexpression de la protéine native ne crée pas de transformation maligne. Ces études montrent un lien direct entre la forme d'une sous-unité d'eIF3 et la transformation d'une cellule en cellule maligne ([Mayeur and Hershey, 2002](#)). Même si le mode d'action de cette forme de la protéine reste inconnu, il est proposé qu'eIF3e agisse comme une onco-protéine négative dominante et donc un répresseur du développement tumoral. Ceci concorde avec les niveaux anormalement faibles d'eIF3e observés dans des cancers du sein et du poumon ([Marchetti et al., 2001](#)).

Les sous-unités eIF3b et eIF3c sont respectivement anormalement abondantes dans des cancers du sein et des testicules ([Lin et al., 2001](#); [Rothe et al., 2000](#)). En outre, les niveaux d'expression des sous-unités impliquées dans l'oncogenèse ont pu être liés avec le diagnostic de ces cancers. En effet, des patients présentant un niveau plus élevé d'eIF3a dans des cellules cancéreuses du sein et de l'estomac ont un taux de survie globalement supérieur à celui de patients présentant une expression moindre ([Chen and Burger, 1999, 2004](#)). De même, l'expression réduite d'eIF3e dans les cancers du poumon ont pu permettre de prédire un faible taux de survie ([Buttitta et al., 2005](#)). Enfin, une expression élevée d'eIF3h dans des cancers de la prostate est également associée à une faible survie ([Saramaki et al., 2001](#)).

CONCLUSION

A travers la description du système de démarrage eucaryotique, nous avons pu percevoir la complexité de ce mécanisme et quelques exemples de sa régulation. La place centrale qu'occupe la traduction justifie un réseau aussi dense d'interconnexions afin de permettre un contrôle très ajusté de cette machinerie. Encore plus au cœur de cette machinerie, le ribosome et ses caractéristiques permettent d'illustrer la possibilité d'un monde tout ARN où des ribozymes auraient peu à peu acquis stabilité et diversité. La traduction des ARNm sans séquence promotrice sur des ribosomes 70S sans l'intervention de facteurs protéiques est un bel exemple soutenant cette thèse. La plongée dans ce monde laisse planer des interrogations. Quand l'ADN serait-il apparu ? A quel moment les protéines sont-elles devenues nécessaires ? Comment cette machinerie de la traduction s'est-elle mise en place ? Autant de questions passionnantes qui demanderont beaucoup de temps et d'efforts de recherche.

De plus, la mise en parallèle du démarrage eucaryotique avec les systèmes bactériens et archéens est tout aussi étonnante tant on aperçoit l'ombre de l'ancêtre universel commun à tous les organismes vivants. Beaucoup d'éléments restent à apporter à ce sujet pour confirmer les prédictions mais la composition de sa machinerie traductionnelle semble précisée. On imagine un système primitif démarrant à partir d'un ribosome 70S, l'activité antiassociatrice du facteur IF3/e/aIF1 serait apparue plus tard pour améliorer l'efficacité du processus. Les facteurs universels IF1/e/aIF1A et IF2/e/aIF5B occuperaient le site A afin de diriger le démarrage de la traduction vers le site P. Une fonction de ces facteurs primitifs reste notamment à préciser : un facteur lie-t-il l'ARNt initiateur ? La réponse à cette question permettrait de savoir si l'ajustement de l'ARNt initiateur a toujours été un prérequis ou bien si cette fonction est apparue pour implémenter un processus peu efficace.

L'étude des systèmes de démarrage actuels est notre seule base pour répondre à ces questions. On perçoit assez simplement la plupart des avantages acquis au cours de l'évolution. La traduction est mieux contrôlée, plus efficace et plus fidèle. Mais la description de ces mécanismes laisse entrevoir le travail encore très important à fournir pour les caractériser parfaitement. Notamment, la récente implication d'e/aIF5B, en plus d'e/aIF2, dans les étapes finales du démarrage laisse quelques zones d'ombre quant aux rôles précis de ce facteur. Les études que nous avons menées afin de clarifier cette situation sont présentées dans le chapitre 2.

La description du facteur eIF3 et de son implication fonctionnelle révèle un système complexe. Son activité repose sur tout un réseau d'interactions, et la modulation de ce réseau confère à eIF3 un rôle critique à bien des égards. Notamment, il apparaît comme un facteur clé pour la traduction de certaines classes d'ARNm par modification de sa composition. Enfin, de plus en plus d'études impliquent ses sous-unités dans les mécanismes de la croissance cellulaire et de l'oncogenèse. Ces sous-unités peuvent ainsi représenter des cibles thérapeutiques importantes pour le traitement de cancers et leurs diagnostics. Toutes ces pistes concordent et nourrissent l'intérêt de la communauté scientifique à étudier ce facteur de démarrage. Néanmoins, la grande complexité du facteur humain limite considérablement son étude. La plupart des équipes purifie le facteur à partir d'extraits bruts de cellules produisant une sous-unité fusionnée à une étiquette. La structure du domaine RRM d'eIF3b ouvre quant à elle la voie à la compréhension des interactions entre les différents partenaires ([ElAntak et al., 2007](#)). La structure obtenue plus tôt de la plus petite sous-unité humaine eIF3k par cristallographie et diffraction aux rayons X n'a pas pu apporter beaucoup de renseignements en raison de la faible caractérisation biochimique de cette protéine ([Wei et al., 2004](#)). Bien qu'étant considérablement simplifié, comparé au facteur humain, le complexe eIF3 de la levure *S. cerevisiae* a permis d'obtenir les premiers éléments quant à son organisation. Les connaissances reposent majoritairement sur des données biochimiques et cellulaires. La résolution de structures tridimensionnelles de ces protéines reste essentielle pour comprendre la manière dont eIF3 orchestre l'ensemble des interactions moléculaires nécessaires au démarrage. Le chapitre 3 présente les résultats que nous avons obtenus sur le facteur eIF3 de *S. cerevisiae*.

Chapitre II Le facteur e/aIF5B

A mon arrivée au laboratoire, l'homologie fonctionnelle entre le facteur de démarrage bactérien IF2, d'une part, et les facteurs eucaryotique et archéen e/aIF5B n'était que partielle. Leurs interventions dans l'étape d'association des sous-unités ribosomales nécessitant leurs activités GTPasique a permis de lier ces facteurs qui partagent une forte homologie de séquence ([Antoun et al., 2003](#); [Grigoriadou et al., 2007a](#); [Shin et al., 2002](#)).

Dans le cas d'IF2, une liaison du fMet-ARNt_f^{Met} par le facteur a pu être mise en évidence en solution ([Petersen et al., 1979](#); [Sundari et al., 1976](#)). De plus, le domaine C-terminal en tonneau β est suffisant pour la liaison du fMet-ARNt_f^{Met} ([Guenneuges et al., 2000](#); [Spurio et al., 2000](#)). Enfin, l'état nucléotidique du facteur IF2 n'influence pas sa liaison au fMet-ARNt_f^{Met} ([Petersen et al., 1979](#)). Ce mode d'interaction a été confirmée par la structure du facteur et de son ligand sur le ribosome en cryomicroscopie électronique ([Allen et al., 2005](#); [Myasnikov et al., 2005](#)).

En revanche, aucune interaction avec l'ARNt initiateur n'avait pu être montrée pour son homologue eucaryotique/archéen e/aIF5B malgré des tentatives répétées ([Roll-Mecak et al., 2000](#)). Les tests d'interaction basés sur la rétention de complexe sur nitrocellulose généralement utilisés sont hors équilibre et donc peu sensibles en cas d'interaction faible. Nous avons donc voulu regarder si une interaction entre e/aIF5B et l'ARNt initiateur pouvait être mise en évidence en utilisant des méthodes à l'équilibre, potentiellement plus sensibles.

PARTIE A CLONAGE ET PURIFICATION DES FACTEURS DE DEMARRAGE E/AIF5B ET AIF1A

I Clonage

A mon arrivée au laboratoire, les gènes codant pour les facteurs archéens avaient été clonés dans des vecteurs pET3alpa permettant de produire les protéines chez *E. coli*. Le vecteur pET3alpa est dérivé du plasmide commercial pET3a (Novagen) de la même façon qu'a pu être obtenu le vecteur pET15blpa (**figure 3.22**). A la différence du vecteur pET15b, ce plasmide permet de produire les protéines sous forme native, sans être fusionnée à une étiquette. Le plasmide permettant de produire une forme comportant les trois domaines N-terminaux du facteur de *S. solfataricus* a été obtenu par mutagenèse dirigée en introduisant un codon stop (TGA) à la place du codon GAA en position 455. Le vecteur d'expression du domaine IV isolé a été construit par amplification PCR à partir du plasmide produisant la protéine native en utilisant des oligonucléotides permettant d'insérer un site *NdeI* et donc un codon ATG à la place du codon du résidu 462. L'insert a ensuite été cloné entre les sites de restriction *NdeI* et *XhoI* du plasmide pET3alpa ([voir Matériels et Méthodes](#)).

Une interaction entre les facteurs eIF1A et eIF5B a pu être mise en évidence en solution ([Marintchev et al., 2003](#)). Deux plasmides de type pET3alpa permettant d'exprimer le facteur aIF1A de *Sulfolobus solfataricus*, d'une part, et celui de *Pyrococcus abyssi*, d'autre part, étaient disponibles au laboratoire. Nous avons ainsi pu tester l'influence de cette interaction sur l'effet d'eIF5B dans le test.

Le facteur eucaryotique présente, par rapport aux facteurs d'Archées, une longue extension N-terminale de 400 résidus chez la levure *Saccharomyces cerevisiae* et de 600 résidus pour le facteur humain. Il a cependant pu être montré que le facteur archéen aIF5B peut pallier partiellement *in vitro* et *in vivo* le défaut provoqué par l'inactivation du gène *FUN12* codant pour eIF5B dans une souche de levure ([Lee et al., 1999](#)). En outre, le facteur humain ou le facteur de levure tronqués de leurs extensions N-terminales permettent d'assurer la croissance de souches dénuées du facteur eIF5B ([Lee et al., 1999](#)). Ces résultats indiquent que les protéines archéennes ou eucaryotiques tronquées du domaine N-terminal doivent agir de façon similaire.

D'autre part, la troncature en N-terminal nous a permis d'exprimer un facteur eIF5B eucaryotique soluble et de le purifier (γ - Δ eIF5B). En effet, des essais de surproduction menés sur le gène codant pour le facteur eIF5B entier dans un vecteur de type pET15blpa ont montré que la protéine est produite sous forme insoluble alors que la forme tronquée est exprimée sous forme soluble.

Le clonage du gène codant cette forme tronquée a été réalisé par l'insertion dans le vecteur pET15blpa du produit de l'amplification de l'ADN génomique de la souche W303a par des oligonucléotides introduisant un codon de démarrage ATG au niveau du codon de la sérine 395, ainsi qu'un site *NdeI*, en amont et un site *BamHI* en aval du gène.

II Surproduction et purification

La souche Rosetta™(DE3)-pLacI RARE a été utilisée pour exprimer les différentes protéines. L'expression et la purification des différents facteurs sont détaillées dans la partie Matériels et Méthodes. Brièvement, la purification des protéines archéennes met à profit la thermostabilité des protéines de ces organismes hyperthermophiles. Après la sonication, le surnageant est chauffé à 70°C pendant 10 minutes afin de précipiter la plupart des protéines d'*E. coli*. Le surnageant de chauffage est ensuite précipité au sulfate de streptomycine afin d'éliminer par centrifugation une partie des acides nucléiques. Les protéines sont alors concentrées par une précipitation au sulfate d'ammonium avant d'être appliquées sur une résine échangeuse d'anions à basse force ionique ([voir Matériels et Méthodes](#)). Le facteur aIF5B ayant un point isoélectrique basique, il passe au travers de cette colonne alors que les acides nucléiques restent fixés sur la résine. La protéine récupérée peut ensuite être accrochée sur une résine échangeuse de cations afin de la séparer d'autres contaminants. Un passage sur un tamis moléculaire permet d'affiner la purification (**Figure 2.1**).

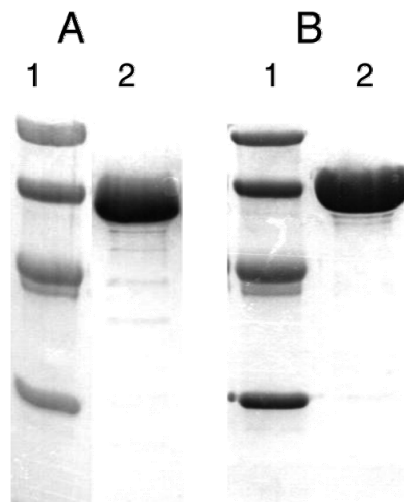


Figure 2.1 : Analyse SDS-PAGE (12%) des protéines ss-aIF5B (A-piste 2) ou γ - Δ eIF5B (B-piste 2). Pistes 1 : Marqueur de poids moléculaires LMW (94, 67, 43 et 30kDa, GE Healthcare)

Le facteur aIF1A de *S. solfataricus* a été purifié en utilisant un protocole similaire excepté pour l'étape de précipitation au sulfate de streptomycine ([voir Matériels et Méthodes](#)).

Les caractéristiques des protéines de levure ne permettent pas de réaliser l'étape de chauffage de l'extrait brut. Pour faciliter la purification de la protéine, le gène codant pour le facteur eIF5B a été cloné dans un vecteur de type pET15b permettant de produire la protéine fusionnée à une étiquette His₆. Ainsi, après la sonication, le surnageant est appliqué sur une résine TALON puis élué avant d'être appliqué sur une résine échangeuse de cations. Le point isoélectrique de cette protéine permet son accrochage sur ce type de résine et la séparation de contaminants par application d'un gradient de sel adapté. La pureté de l'échantillon est alors affinée par un passage sur un tamis moléculaire ([voir Matériels et Méthodes](#)) (**Figure 2.1**).

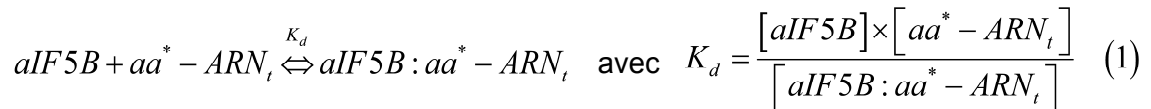
PARTIE B PROTECTION DU MET-ARNt INITIATEUR PAR E/aIF5B

I Le test de protection de la déacylation spontanée d'un ARNt aminoacylé

1 Principe du test

Chez les Bactéries, la liaison du fMet-ARNt_f^{Met} à IF2 conduit protéger l'acide aminé estérifié en 3' de l'ARNt de la déacylation spontanée (Petersen et al., 1979). Nous avons supposé qu'une interaction entre e/aIF5B et l'ARNt initiateur aminoacylé pourrait être détectée en utilisant cette même technique. Le test de protection tire parti de la faible stabilité de la liaison ester qui lie l'acide aminé à l'extrémité 3' de l'ARNt. Or, si une protéine lie l'ARNt par la tige acceptrice, l'attaque de l'eau sur la fonction ester est empêchée, ce qui fait diminuer la vitesse d'hydrolyse. Afin de suivre la déacylation de l'ARNt, l'acide aminé qui y est aminoacylé est marqué radioactivement (aa*).

Ainsi, si l'on considère que la formation du complexe entre le facteur aIF5B et l'ARNt aminoacylé est rapide devant la vitesse de déacylation, il peut être écrit à tout instant :



Avec :

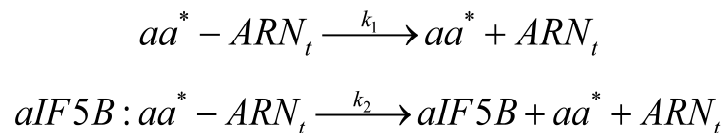
Kd la constante de dissociation du complexe aIF5B:aa*-ARNt

[aIF5B] la concentration de facteur libre

[aa*-ARNt] la concentration d'ARNt aminoacylé libre

[aIF5B:aa*-ARNt] la concentration de complexe à l'état stationnaire

Les vitesses de déacylation de l'ARNt aminoacylé libre ou lié au facteur sont différentes :



La vitesse de déacylation, v, s'écrit :

$$v = \frac{d[ARN_t]}{dt} = k_1 \times [aa^* - ARN_t] + k_2 \times [aIF5B : aa^* - ARN_t] \quad (4)$$

Si on note [aIF5B]₀ la concentration totale en facteur et [aa*-ARNt]₀ la concentration en ARNt aminoacylé ajouté au temps 0, alors la loi de conservation des espèces s'écrit :

$$[aIF5B]_0 = [aIF5B] + [aIF5B : aa^* - ARN_t] \quad (2)$$

$$[aa^* - ARN_t]_0 = [aa^* - ARN_t] + [ARN_t] + [aIF5B : aa^* - ARN_t] \quad (3)$$

Où [ARNt] est la concentration d'ARNt déacylé.

Dans des conditions où $[aa^*-ARNt]_0 \ll K_d$, peu de facteur est engagé dans la liaison de l'ARNt et alors $[aIF5B:aa^*-ARNt] \ll [aIF5B]$ et donc $[aIF5B]_0 \simeq [aIF5B]$. La relation (1) peut également s'écrire :

$$[aIF5B : aa^* - ARN_t] \simeq \frac{[aIF5B]_0 \times [aa^* - ARN_t]}{K_d} \quad (5)$$

En injectant la relation (5) dans l'équation (4), on obtient :

$$v \simeq \left(k_1 + k_2 \times \frac{[aIF5B]_0}{K_d} \right) \times [aa^* - ARN_t] \quad (6)$$

De même, on peut réécrire la relation (3) :

$$[aa^* - ARN_t] \simeq \left(\frac{K_d}{K_d + [aIF5B]_0} \right) \times \left([aa^* - ARN_t]_0 - [ARN_t] \right) \quad (7)$$

En substituant l'équation (7) dans la relation (6), on obtient finalement :

$$v = \frac{d[ARN_t]}{dt} \simeq \left[k_1 + (k_2 - k_1) \times \left(\frac{[aIF5B]_0}{K_d + [aIF5B]_0} \right) \right] \times \left([aa^* - ARN_t]_0 - [ARN_t] \right)$$

Ainsi, si l'on intègre cette équation en considérant que lorsque $t \rightarrow 0$, $[ARN_t] \rightarrow 0$, puisque au temps $t=0$, on considère que tout l'ARNt est aminoacylé, la déacylation de l'ARNt suit une loi exponentielle :

$$A(t) = \left([aa^* - ARN_t]_0 - [ARN_t] \right) \simeq [aa^* - ARN_t]_0 \times e^{-B \times t}$$

$$\text{avec } B = \left[k_1 + (k_2 - k_1) \times \left(\frac{[aIF5B]_0}{K_d + [aIF5B]_0} \right) \right] \quad (8)$$

La quantité d'ARNt aminoacylé présent à l'instant t (qui correspond à la grandeur $A(t)$) peut être mesurée par précipitation de l'ARNt à l'acide trichloroacétique et mesure de la radioactivité retenue sur le filtre. Les comptages à différents temps décrivent ainsi une exponentielle dont la constante de temps peut-être évaluée par minimisation de l'écart entre la courbe expérimentale et la courbe exponentielle théorique (programme MC-Fit, [Dardel, 1994](#)). Les constantes de vitesse déduites des constantes de temps sont mesurées pour différentes valeurs de concentration en facteur afin de tracer la courbe $B=f([aIF5B]_0)$ permettant ainsi d'évaluer la valeur de la constante de dissociation du complexe K_d par ajustement avec la courbe théorique (programme MC-Fit).

2 Mise en œuvre expérimentale

La vitesse de déacylation est sensible aux conditions réactionnelles, notamment aux concentrations en sels et en glycérol. Le facteur eIF5B de levure est soluble à haute force ionique (500mM NaCl). Ainsi, pour éviter des variations artéfactuelles de la vitesse de déacylation, tous les facteurs purifiés sont conservés dans un tampon sans glycérol et le milieu réactionnel est complété en chlorure de sodium afin d'assurer une concentration constante égale à 500mM pendant les tests de protection du Met-ARNt_i^{Met}, quelle que soit la concentration d'eIF5B présente. Le suivi de la déacylation en présence des facteurs archéens s'effectue à 51°C dans un tampon contenant {20mM HEPES pH8,0 ; 100mM KCl ; 500mM NaCl ; 5mM MgCl₂ ; 1mM DTT ; 0,1mM EDTA ; 0,2mg/mL BSA (Bovin Albumin Serum, Roche) ; 5% glycérol ; 1mM GTP ou GDPNP}. En absence de facteur, le Met-ARNt_i^{Met} est complètement déacylé en 30 minutes à 51°C. L'utilisation du facteur de levure eIF5B ne permet pas de travailler à ces températures élevées et les vitesses de déacylation ont été mesurées à 30°C. Dans ces conditions, la vitesse de déacylation spontanée est de 0,04min⁻¹, une vitesse de déacylation acceptable pour réaliser les expériences (déacylation complète de l'ARNt en 100min.). Dans l'ensemble des tests, la concentration en ARNt aminoacylé utilisée est 80nM. Nous avons vérifié *a posteriori* que cette concentration est largement inférieure aux constantes de dissociation mesurées, permettant ainsi de réaliser les approximations évoquées précédemment. La procédure expérimentale est expliquée plus en détail dans la partie Matériels et Méthodes.

II Résultats

1 Protection de l'ARNt initiateur méthionylé par aIF5B

Nous avons dans un premier temps testé la liaison du Met-ARNt_i^{Met} de *S. solfataricus* au facteur aIF5B de cet organisme (ss-aIF5B). Le gène codant pour cet ARNt initiateur de cet organisme a été obtenu par assemblage d'oligonucléotides. L'ARNt initiateur, cloné dans le vecteur pET3alpa, a été exprimé dans la souche Rosetta™(DE3)pLacIRARE, puis purifié. L'ARNt a finalement été aminoacylé par la méthionyl-ARNt synthétase d'*E. coli*. A 51°C, la vitesse de déacylation de cet ARNt est de 0,14min⁻¹. L'ajout d'une concentration finale de 50µM ss-aIF5B et 1mM GTP permet une réduction de la vitesse à 0,01min⁻¹, indiquant clairement la liaison de l'ARNt par ce facteur. La constante de dissociation finalement déterminée est de 2,7µM ± 0,5µM (**Figure 2.2**). Cette constante de dissociation est du même ordre de grandeur que celle mesurée pour le complexe binaire IF2:fMet-ARNt_f^{Met} ([Petersen et al., 1979](#); [Spurio et al., 2000](#)) mais est considérablement plus élevée que celles mesurées pour le facteur eIF2 ([Nika et al., 2001](#); [Yatime et al., 2004](#)).

De façon similaire, nous avons produit l'ARNt initiateur de *P. abyssi*. Après purification, la protection de cet ARNt initiateur méthionylé par le facteur aIF5B de l'Archée *Pyrococcus abyssi* (pa-aIF5B) a été mesurée. Une constante de dissociation de 1,9µM ± 0,4µM a pu être déterminée.

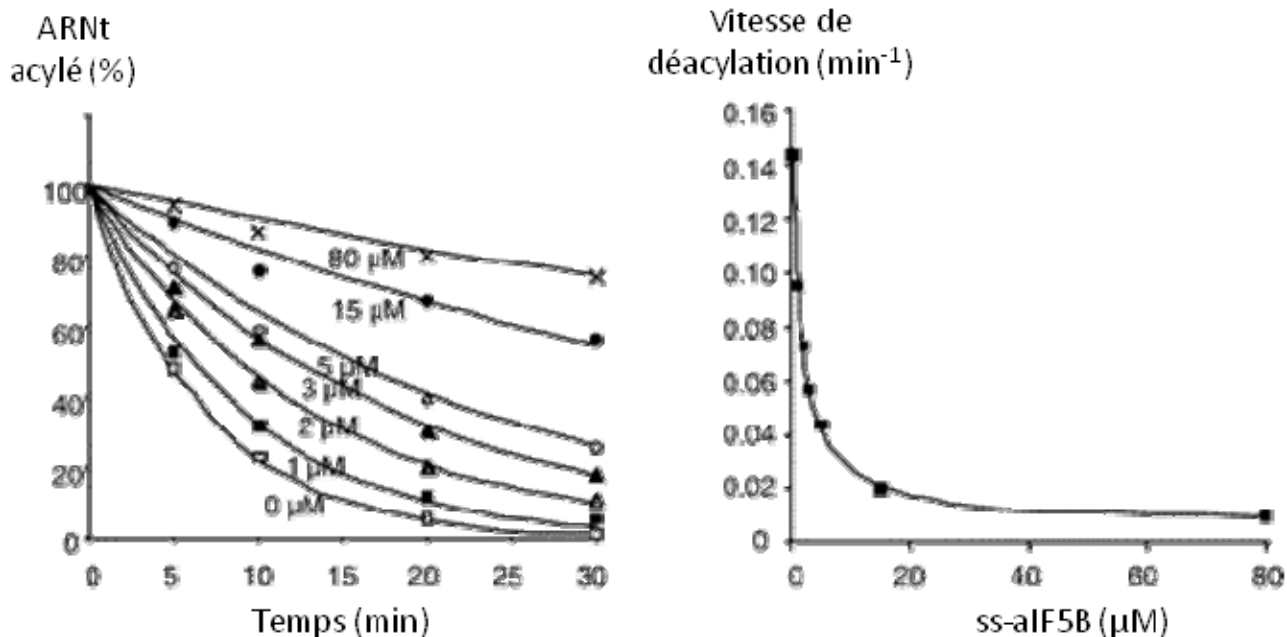


Figure 2.2 : Détermination de la constante de dissociation (K_d) du complexe entre ss-aIF5B et le Met-ARNt_i^{Met} de *S. solfataricus*. La figure de gauche représente les cinétiques de déacylation à différentes concentrations de facteur. Les vitesses de déacylation sont ensuite reportées en fonction de la concentration en facteur (graphique de gauche).

Les ARNt initiateur archéens et bactériens ont des séquences très proches. Ils diffèrent notablement par la paire de bases 1 et 72, systématiquement non appariée chez les Bactéries et présentant un appariement faible (A1:U72) chez les Archées (**Figure 1.3, p8**). Les constantes de dissociation des complexes formés par l'ARNt initiateur d'*E. coli* avec ss-aIF5B ou pa-aIF5B sont très proches de celles mesurées avec leurs ARNt initiateurs naturels (**Tableau 2.1**). Il a donc été adopté dans la suite de l'étude, l'ARNt initiateur d'*E. coli* comme ARNt initiateur modèle.

Protéine	Ligand ARNt	K_d (μM)
ss-aIF5B	Met-ARNt _i ^{Met} <i>S. solfataricus</i>	2.7 ± 0.5
ss-aIF5B	Met-ARNt _f ^{Met} <i>E. coli</i>	4.2 ± 0.7
pa-aIF5B	Met-ARNt _i ^{Met} <i>P. abyssi</i>	1.9 ± 0.4
pa-aIF5B	Met-ARNt _f ^{Met} <i>E. coli</i>	2.7 ± 0.7

Tableau 2.1 : Influence de la nature de l'ARNt initiateur sur sa liaison par les facteurs ss-aIF5B et pa-aIF5B.

L'influence de l'état nucléotidique du facteur a ensuite été testée en remplaçant dans le milieu réactionnel le GTP par du GDPNP ou du GDP. Dans les deux cas, les constantes de dissociations mesurées sont très proches de celles obtenues en présence de GTP (**Tableau 2.2**). Comme dans le cas d'IF2 bactérien, la liaison de l'ARNt initiateur méthionylé à e/alf5B ne dépend pas de l'état nucléotidique du facteur ([Grigoriadou et al., 2007a](#); [Petersen et al., 1979](#)).

Protéine	Ligand ARNt	Nucléotide (1mM)	K _d (μM)
ss-alf5B	Met-ARNt _f ^{Met}	GTP	4.2 ± 0.7
ss-alf5B	Met-ARNt _f ^{Met}	GDP	3.5 ± 0.7
ss-alf5B	Met-ARNt _f ^{Met}	GDPNP	3.9 ± 0.7

Tableau 2.2 : Influence du nucléotide lié à ss-alf5B sur sa liaison à l'ARNt initiateur d'*E. coli*.

2 Compétition par l'ARNt_f^{Met}

Nous avons observé que l'ARNt initiateur non aminoacylé n'est pas un inhibiteur compétitif de la liaison de l'ARNt méthionylé à alf5B. En effet, en condition équimolaire d'ARNt_f^{Met} et de Met-ARNt_f^{Met}, en présence d'alf5B, les vitesses de déacylation mesurées sont inchangées. Aucun effet n'est observé si l'ARNt_f^{Met} est ajouté en large excès (10μM). Ainsi, l'aminoacylation de l'ARNt initiateur est requise pour sa liaison à alf5B.

3 Spécificité de la liaison

Le déterminant reconnu par le facteur de démarrage a pu être identifié par l'utilisation de divers ARNt représentant des ligands potentiels. Dans un premier temps, la protection de deux ARNt élongateurs lysine et valine d'*E. coli* (Lys-ARNt^{Lys} et Val-ARNt^{Val}) par le facteur ss-alf5B a été testée. Une faible protection du Lys-ARNt^{Lys} a pu être mise en évidence pour la concentration la plus élevée du facteur (80μM). La constante de dissociation déduite de cette expérience est de l'ordre de 50μM (**Figure 2.3**). La vitesse de déacylation du deuxième ligand testé, Val-ARNt^{Val}, est faible dans les conditions du test à 51°C (0,03min⁻¹). Le milieu réactionnel a été adapté en substituant le tampon HEPES par un tampon Bicine-NaOH pH 9,0 à une concentration finale de 100mM. Ceci a permis d'augmenter la vitesse de déacylation du Val-ARNt^{Val} en l'absence de facteur à 0,17min⁻¹. Dans ces conditions, une faible protection est observée en présence d'une concentration de 80μM en facteur (**Figure 2.3**). Il est à souligner que dans ce même tampon à pH 9, l'affinité du Met-ARNt_f^{Met} pour ss-alf5B est identique à celle mesurée à pH 8 (4,2 ± 0,7μM). Nous pouvons donc en conclure que la liaison d'un ARNt aminoacylé à ss-alf5B présente une certaine spécificité. Celle-ci peut être liée soit à la chaîne latérale de la méthionine, soit à des particularités de la séquence de l'ARNt initiateur.

Afin de préciser les éléments reconnus par ss-aIF5B, deux ARNt modifiés génétiquement ont été utilisés. Le premier correspond à l'ARNt initiateur d'*E. coli* dont l'anticodon (CAU) a été muté par en anticodon valine (UAC). Cet ARNt noté $\text{ARNt}_f^{\text{Met}}_{[\text{UAC}]}$ est aminoacylable par la valyl-ARNt synthétase (Schulman and Pelka, 1988). Un second ARNt dérive de l'ARNt élongateur valine d'*E. coli* dont l'anticodon est changé en anticodon méthionine (CAU). A l'inverse, cet ARNt, l' $\text{ARNt}_1^{\text{Val}}_{[\text{CAU}]}$, peut être méthionylé par la méthionyl-ARNt synthétase. Lorsque l' $\text{ARNt}_f^{\text{Met}}_{[\text{UAC}]}$ porte un acide aminé valine, une faible protection de la liaison ester de l'ARNt est détectée (Figure 2.3). La constante de dissociation, estimée à $50\mu\text{M}$, correspond à une diminution d'un ordre de grandeur de l'affinité pour ss-aIF5B par rapport à celle mesurée pour le $\text{Met-ARNt}_f^{\text{Met}}$. A l'inverse, la simple présence de l'acide aminé méthionine sur l' $\text{ARNt}_1^{\text{Val}}$ suffit pour protéger efficacement la liaison ester, grâce à la formation d'un complexe avec ss-aIF5B d'une constante de dissociation de $2,7\mu\text{M} \pm 0,5\mu\text{M}$. De la même façon, l' $\text{ARNt}_m^{\text{Met}}$ élongateur d'*E. coli* à ss-aIF5B aminoacylé par une méthionine se lie à ss-aIF5B avec une constante de dissociation de $3,3\mu\text{M} \pm 0,5\mu\text{M}$. Nous en concluons que la sélectivité de la liaison d'un ARNt aminoacylé par la protéine ss-aIF5B repose essentiellement sur la reconnaissance de la méthionine estérifiée.

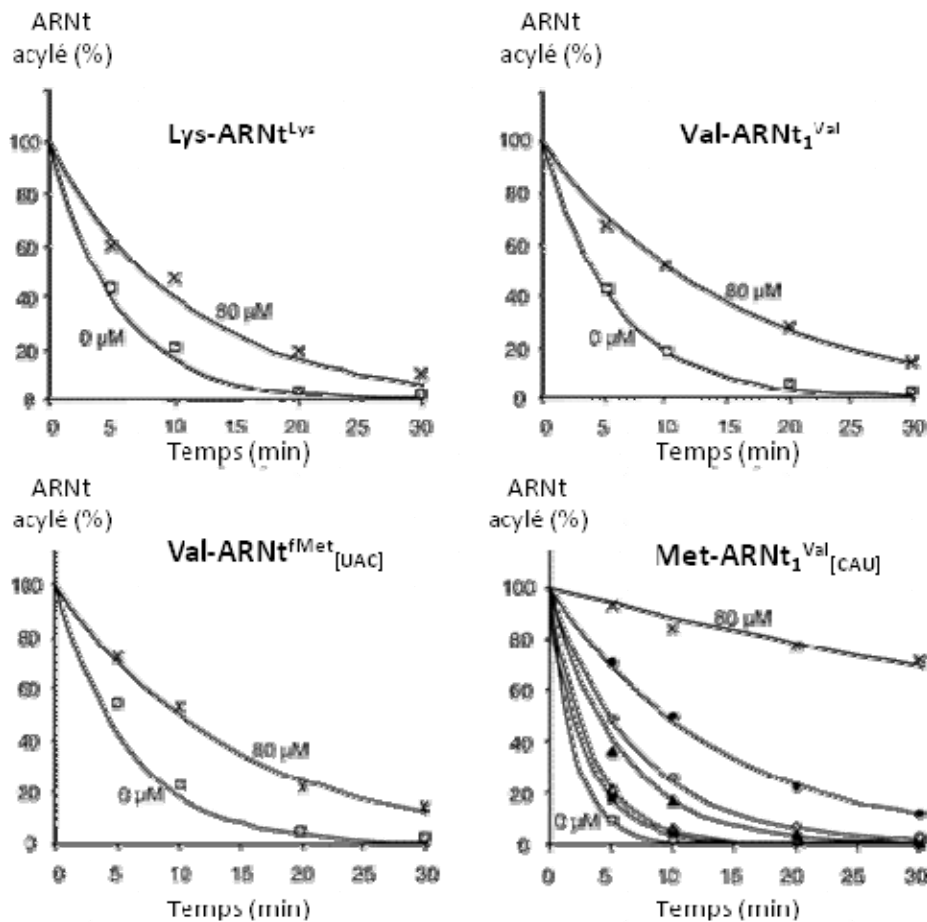


Figure 2.3 : Spécificité de la liaison de l'ARNt aminoacylé par ss-aIF5B. Les trois premiers graphiques représentent les cinétiques de déacylation de l'ARNt aminoacylé indiqué ($\text{Lys-ARNt}^{\text{Lys}}$, $\text{Val-ARNt}_1^{\text{Val}}$, or $\text{Val-ARNt}_f^{\text{Met}}_{[\text{UAC}]}$) en présence de $80\mu\text{M}$ ss-aIF5B (x) ou en l'absence de facteur (\square). Dans le cas de $\text{Met-ARNt}_1^{\text{Val}}_{[\text{CAU}]}$, les courbes de déacylation sont reportées pour plusieurs concentrations de ss-aIF5B [(■) 1, (\triangle) 2, (\blacktriangle) 5, (\circ) 8, and (\bullet) 15 μM] sur le quatrième graphique.

4 Cartographie de l'interaction

Les mesures de protection de la déacylation spontanée du Met-ARNt^{Met} avec la protéine correspondant aux trois domaines C-terminaux de ss-aIF5B indiquent que la délétion du tonneau β C-terminal du facteur abolit sa liaison au Met-ARNt^{Met} (**Tableau 2.3**). A l'inverse le domaine C-terminal isolé assure une liaison du polynucléotide comparable à celle du facteur entier (**Tableau 2.3**). Cette liaison assurée par le domaine C-terminal n'est pas modifiée si le domaine N-terminal isolé y est ajouté. De plus, la spécificité de la liaison réside bien dans ce domaine C-terminal puisque le Val-ARNt^{Val}, le Val-ARNt^{Met}_[UAC] et le Lys-ARNt^{Lys} ne forment pas de complexes de forte affinité (**Figure 2.4**). Il peut alors être conclu que, comme dans le cas d'IF2, le domaine IV C-terminal du facteur aIF5B est seul responsable de la liaison de l'ARNt aminoacylé.

Protéine	Ligand ARNt	Kd (μM)
ss-aIF5B	Met-ARNt ^{Met}	4.2 ± 0.7
Domaine IV de ss-aIF5B	Met-ARNt ^{Met}	1.6 ± 0.3
Domaine I-II-III de ss-aIF5B	Met-ARNt ^{Met}	>80

Tableau 2.3 : Liaison du Met-ARNt^{Met} aux domaines de ss-aIF5B

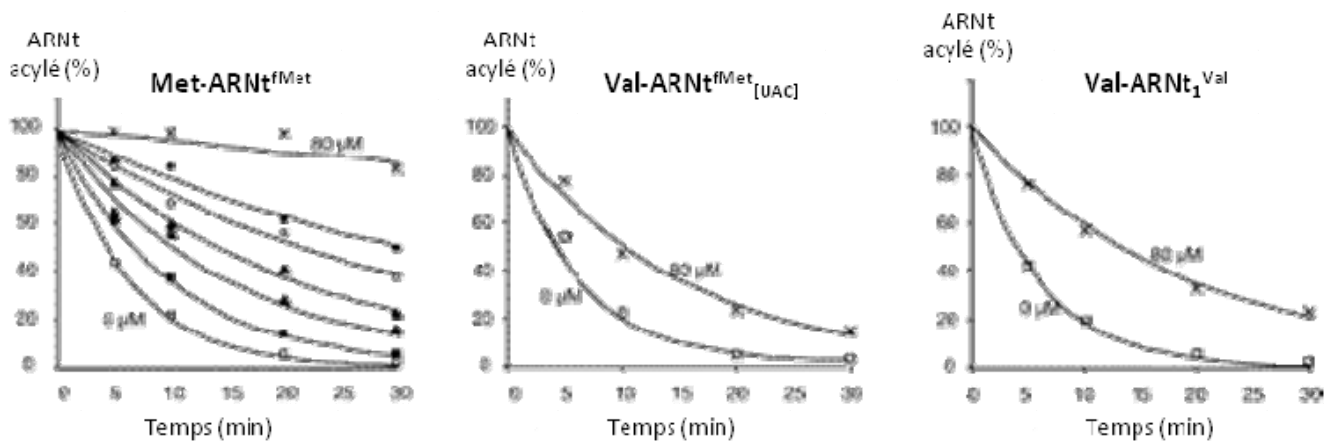


Figure 2.4 : Spécificité de la liaison de l'ARNt aminoacylé par le domaine IV de ss-aIF5B. Les trois graphiques représentent les cinétiques de déacylation de l'ARNt aminoacylé indiqué en présence de 80 μM ss-aIF5B (x) ou en l'absence de facteur (□). Dans le cas du Met-ARNt^{Met}, les courbes de déacylation sont reportées pour plusieurs concentrations du domaine IV de ss-aIF5B [(■) 1, (▲) 2, (▲) 5, (○) 8, and (●) 15 μM].

5 Influence de la liaison à aIF1A

Dans un premier temps, nous avons tenté de mettre en évidence l'interaction entre aIF1A et aIF5B par des expériences de retard sur gel. Ces protéines ont des points isoélectriques basiques, nous avons donc utilisé des gels natifs de pH acide. Différents ratios des protéines aIF5B de *S. solfataricus*, d'une part, et de *P. abyssi*, d'autre part, ont été testés. Cependant, il semblerait que l'interaction soit trop faible pour pouvoir être observée par cette technique. La migration des protéines individuelles est bien visible.

Dans un deuxième temps, nous avons analysé l'influence de la présence d'aIF1A sur la constante de dissociation du complexe aIF5B:Met-ARNt. Les mesures de protection de la déacylation spontanée du Met-ARNt^{Met} ont été réalisées avec différents rapports molaires d'aIF5B et d'aIF1A. Un excès stœchiométrique d'un facteur 2 de pa-aIF1A par rapport à pa-aIF5B (1µM) a été testé. Dans ce cas, les vitesses de déacylation évaluées sont très proches (0,074min⁻¹ pour pa-aIF5B seul et 0,068min⁻¹ en présence de pa-aIF5B et pa-aIF1A). Des résultats similaires ont été obtenus avec les protéines de *S. solfataricus*. Ces résultats suggèrent que le facteur aIF1A n'influe pas sur la liaison du Met-ARNt^{Met} par le facteur aIF5B en solution.

6 Protection de l'ARNt initiateur méthionylé par eIF5B

Les résultats obtenus dans le cas des facteurs archéens nous ont incités à les étendre au système eucaryotique. Nous avons produit le facteur eIF5B de *S. cerevisiae* sous une forme dénuée de l'extension N-terminale caractéristique des Eucaryotes. Cette forme raccourcie semble assurer les fonctions essentielles du facteur entier. Cette protéine notée γ - Δ eIF5B a été produite sous forme fusionnée à une étiquette pour faciliter sa purification. Des concentrations élevées de ce facteur permettent de protéger le Met-ARNt^{Met} d'*E. coli* de sa déacylation spontanée à 30°C. Ainsi, pour la concentration la plus élevée en facteur dans ce test (122µM), la vitesse de déacylation est réduite à 0,01 min⁻¹ alors qu'un milieu réactionnel sans facteur permet de déacyler le Met-ARNt^{Met} à une vitesse de 0,04 min⁻¹. Afin de confirmer que l'effet est bien imputable à la protéine eIF5B, cette dernière a été substituée dans un contrôle par 80µM de BSA. Cette protéine inerte n'a qu'un faible impact sur la déacylation qui s'opère à une vitesse de 0,036 min⁻¹. Les mesures de vitesses de déacylation en présence de différentes concentrations de γ - Δ eIF5B nous ont permis d'évaluer une constante de dissociation du complexe de 40µM \pm 5µM à 30°C. La température ne semble pas influencer de façon importante cette valeur puisque la même série d'expérience menée à 25°C permet de déterminer une constante de dissociation indistinguable de 39µM \pm 4µM.

A ce stade, nous ne pouvons exclure que la faible affinité observée soit liée à l'utilisation de l'ARNt^{Met} d'*E. coli* comme ligand modèle. En effet, l'ARNt initiateur de la levure possède plusieurs modifications post-transcriptionnelles et notamment la base 64 est spécifiquement modifiée par une phosphoribosylation ([Desgres et al., 1989](#); [Kiesewetter et al., 1990](#)). Cet ARNt, isolé d'un extrait brut de *S. cerevisiae*, nous a été aimablement fourni par le Dr. Gérard Keith (IBMC, Strasbourg) et a ainsi pu être testé. La constante de dissociation du complexe formé entre γ - Δ eIF5B et l'ARNt^{Met} méthionylé de levure (41µM \pm 10µM, à 30°C) est identique à celle du complexe γ - Δ eIF5B:Met-ARNt^{Met}.

La spécificité de la liaison par le facteur eucaryotique a ensuite été évaluée en utilisant les ARNt élongateurs lysine et valine d'*E. coli*. Une protection pour ces deux ARNt aminoacylés a pu être observée et des constantes de dissociation de l'ordre de 100-150 µM en sont déduites. Enfin, nous avons mesuré une constante de dissociation de 46µM \pm 10µM pour le complexe γ - Δ eIF5B:Met-ARNt₁^{Val}_[CAU]. Cette affinité est similaire à celle observée pour l'ARNt initiateur. Ceci indique que la liaison d'un ARNt par le facteur γ - Δ eIF5B est indépendante de sa séquence nucléotidique. Bien que l'effet soit moins prononcé que dans le cas archéen, il apparaît qu'une méthionine portée par un ARNt soit l'élément permettant au facteur de reconnaître son ligand.

PARTIE C CRISTALLISATION DU FACTEUR

I Obtention des cristaux

Une seule structure tridimensionnelle d'un facteur e/αF5B, celle d'αF5B de *M. thermoautotrophicum* est disponible ([Roll-Mecak et al., 2000](#)). La protéine a été cristallisée en l'absence de nucléotide, et les liaisons du GDPNP et du GDP ont été étudiées par trempage des cristaux, montrant des mouvements relativement faibles du domaine C-terminal en fonction du nucléotide. Dans ces cristaux, la conformation d'αF5B diffère notablement de celle observée par microscopie électronique pour le facteur bactérien IF2 lié au ribosome ([Allen et al., 2005](#)).

Il nous a donc paru intéressant de tenter la cristallisation d'αF5B de *S. solfataricus*, en particulier en présence des différents nucléotides. 3 conditions de cristallisations de la matrice Index (Hampton Research) ont permis d'obtenir des cristaux de la protéine à partir d'une liqueur mère à 3,5 mg/mL et 250 mM NaCl à 24°C (condition 1 : [0,2M Li₂SO₄ ; 0,1M bis-Tris pH 6,5 ; 25% w/v PEG3350] ; condition 2 [0,2M Li₂SO₄ ; 0,1M HEPES pH 7,5 ; 25% w/v PEG3350] ; condition 3 [0,3M (NH₄)₂SO₄ ; 0,1M Tris-HCl pH 8,5 ; 30% w/v PEG3350]). La première condition n'a pu être plus optimisée et seuls des cristaux très petits et mal formés ont été produits. La seconde condition fournissait des cristaux mieux structurés mais de taille insuffisante pour envisager de les faire diffracter. Enfin, la troisième condition identifiée deux mois après les précédentes a permis d'obtenir d'emblée des cristaux bien formés et de taille suffisante pour être testés au synchrotron (environ 50x50x10µm). Les tentatives d'optimisation de la condition de cristallisation sont demeurées infructueuses.

Des cristaux d'une taille plus importante (220x50x10µm) ont cependant pu être obtenus en appliquant 4 µL de la condition de cristallisation en trait sur une lamelle siliconée puis en croisant avec un trait perpendiculaire de 4 µL de la liqueur mère. Le gradient de concentration en agent précipitant ainsi formé permet de moduler l'établissement de l'équilibre des concentrations.

Une seconde approche a consisté à mettre en pratique la méthode en gel. De l'agarose à bas point de solidification (0,1% w/v) est mélangé à l'agent précipitant pour obtenir la concentration désirée. Après ébullition puis refroidissement, la solution est mélangée à la protéine en ratio 1:1 puis l'ensemble est coulé dans un capillaire, qui sera ensuite scellé. Alternativement, l'agarose contenant la condition de cristallisation est réparti directement dans les capillaires et une fois le milieu gélifié, la liqueur mère est apposée au dessus du capillaire puis l'ensemble est scellé. La diffusion de la protéine dans le gel permet ainsi l'établissement d'un gradient pouvant aboutir à la formation de cristaux. Cette technique est censée faciliter la croissance des cristaux dans les trois dimensions, en limitant les phénomènes de convection et de sédimentation des cristaux ([Miller et al., 1992](#)). Toutefois, dans ce cas, nous n'avons obtenu que des cristaux de petites tailles.

Le contenu des cristaux a été testé par dissolution. Les résultats indiquent sans ambiguïté la présence de la protéine αF5B intacte dans ces cristaux.

Des tentatives de cocrystallisation du facteur avec 1mM GTP-Mg²⁺, 1mM GDPNP-Mg²⁺ ou 1mM GDP-Mg²⁺ ont permis de retenir 3 conditions uniques au GDP, 2 uniques au GTP et 2 conditions communes. Néanmoins, les cristaux très petits n'ont pu être optimisés, que ce soit en variant la concentration d'agents précipitants, ou en utilisant d'autres techniques de cristallisation.

II Diffraction

Les cristaux de la protéine aIF5B seule ont été préalablement transférés dans une solution cryprotectante. Sur les trois agents testés, le glycérol a causé la dissolution des cristaux alors que le PEG3350 et le PEG400 n'ont pas affecté la morphologie du cristal. Les boucles sont ensuite refroidies dans l'éthane liquide. Des clichés de diffraction de cinq cristaux ont été enregistrés sur la ligne ID14-eh1 à l'ESRF de Grenoble. Dans l'ensemble les pouvoirs diffractants sont médiocres. Seuls deux cristaux ont permis de déduire une maille cristallographique, indiquant une symétrie P4 et des paramètres de mailles de $169 \times 169 \times 137 \text{Å}^3$. L'unité asymétrique contiendrait ainsi 2 molécules. La meilleure diffraction a mené à l'observation de tâches jusqu'à 6Å de résolution. L'impossibilité d'obtenir des cristaux de meilleure qualité nous a contraints à arrêter ce travail.

DISCUSSION

Notre étude a permis de mettre en évidence, pour la première fois, l'existence d'un complexe e/aIF5B:Met-ARNt^{Met} en solution. La spécificité de la liaison est très similaire à celle observée entre le facteur IF2 bactérien et le fMet-ARNt^{Met} puisque les deux protéines sélectionnent l'acide aminé méthionine ([Li et al., 1996](#); [Petersen et al., 1979](#); [Sundari et al., 1976](#)).

L'analogie entre e/aIF5B et IF2 est renforcée par le fait que, dans les deux cas, le domaine C-terminal en tonneau β (**Figure 2.5, domaine IV**) lie l'ARNt initiateur avec une affinité et une spécificités égales à celles de la protéine entière ([Guenneuges et al., 2000](#); [Spurio et al., 2000](#)). La valeur de la constante de dissociation mesurée, de l'ordre du μM , est de plus inchangée par l'état nucléotidique du facteur. Tout ceci renforce l'idée selon laquelle IF2 et e/aIF5B ont conservé des fonctions communes au cours de l'évolution ([Choi et al., 1998](#); [Guenneuges et al., 2000](#); [Kyrpides and Woese, 1998b](#); [Spurio et al., 2000](#)). Une de ces fonctions communes serait de vérifier la présence d'un ARNt initiateur au site P et d'adapter son positionnement, avant que l'étape d'assemblage des sous-unités ne puisse avoir lieu ([Allen et al., 2005](#); [Grigoriadou et al., 2007a](#); [Lee et al., 2002](#); [Shin et al., 2002](#)).

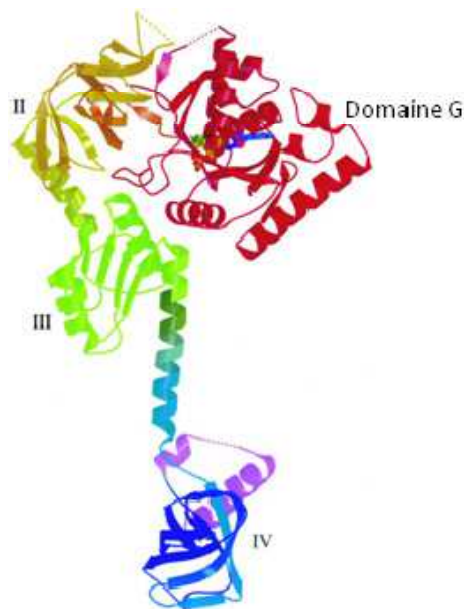


Figure 2.5 : Structure du facteur aIF5B de *Methanobacterium thermoautotrophicum* d'après ([Roll-Mecak et al., 2000](#)). Le code couleur est comme suit : domaine G, rouge ; domaine II, jaune ; domaine III, vert, domaine IV, bleu ; hélices C-terminales, violet.

En revanche, la discrimination semble être moins importante vis-à-vis de l'acide aminé comparée à la sélection de la méthionine N-bloquée opérée par le facteur IF2. En outre, le facteur aIF5B semble être indifférent à la séquence nucléotidique de l'ARNt. Dans le cas eucaryotique que nous avons étudié, la spécificité envers la méthionine est encore plus réduite. Une telle spécificité relativement faible est cohérente avec l'idée que la sélection initiale hautement spécifique de l'ARNt initiateur est gouvernée par le facteur e/aIF2 chez les Archées et les Eucaryotes. La présence des facteurs additionnels dans ces systèmes semble indiquer un rôle de la protéine aux étapes ultimes du démarrage de la traduction et directement sur le ribosome. Ainsi, e/aIF5B pourrait être recruté lorsqu'un ARNt est présent au site P ([Lee et al., 2002](#)). Le recrutement du facteur au ribosome a probablement lieu par l'intermédiaire du facteur e/aIF1A ([Marintchev et al., 2003](#)) et des résultats récents indiquent que cette liaison est importante pour les étapes finales d'association des sous-unités ribosomales ([Acker et al., 2006](#)). Nous ne pouvons cependant pas exclure que cette interaction, ou une interaction du type IF1:IF2 (**voir p.50**), puisse influencer la liaison d'e/aIF5B au Met-ARNt^{Met} sur le ribosome.

Enfin, la faible interaction observée avec le facteur de *S. cerevisiae* est en accord avec l'observation que la délétion de ce gène n'induit qu'un phénotype de croissance lente ([Choi et al., 1998](#)) alors qu'eIF2 est strictement essentiel pour la viabilité de la cellule ([Hannig et al., 1993](#)). La présence d'un jeu réduit de facteurs de démarrages chez les Archées pourrait entraîner une sélection primaire de l'ARNt initiateur moins rigoureuse par rapport aux Eucaryotes. Ainsi, au cours de l'évolution, le facteur archéen aIF5B se serait vu doté d'une spécificité accrue envers un ARNt initiateur lié au site P, par rapport à eIF5B, afin de pallier le nombre réduit de facteurs assistant le ribosome dans la sélection du bon cadre de lecture.

Finalement, il faut remarquer que e/aIF5B ne lie pas un ARNt_f^{Met} méthionylé et formylé (résultats non publiés obtenus au laboratoire). En outre, il faut noter que la traduction bactérienne peut se dérouler sans formylation de l'ARNt initiateur mais de façon beaucoup moins efficace ([Guillon et al., 1992a](#)). Ces éléments tendent à impliquer que la formylation serait apparue avec la branche phylogénétique bactérienne.

Ces résultats ont fait l'objet d'une publication dans le journal *Biochemistry* ([Guillon et al., 2005](#)) et ont amené à la coécriture d'un chapitre dans *Methods in Enzymology* ([Mechulam et al., 2007](#)).

Chapitre III Le facteur eIF3

PARTIE A STRATEGIE DE PURIFICATION DU COMPLEXE EIF3 DE *SACCHAROMYCES CEREVISIAE*

Le facteur de la levure de boulanger représente une forme simplifiée du facteur multimérique humain. Il comprend les cinq sous-unités qui forment le « cœur conservé » du facteur de démarrage eIF3. A ce jour, seule la structure d'un motif d'environ 100 résidus d'une des sous-unités de cœur, eIF3b, a été résolue (ElAntak et al., 2007). Une structure obtenue par cryomicroscopie électronique à 30Å du facteur humain révèle une forme à laquelle il a été attribué une terminologie anthropomorphique (Siridechadilok et al., 2005). Les protubérances sont ainsi identifiées en tant que tête, bras droit et gauche, pied droit et gauche mais il n'a pu être attribué de densité aux différentes sous-unités (Figure 1.22, p.37). En outre, la majorité des résultats biochimiques concerne le facteur de levure. Une progression cohérente consisterait à résoudre la structure du facteur de cet organisme pour le caractériser de manière plus approfondie et donner ainsi des bases solides pour soutenir les découvertes concernant les facteurs plus complexes d'autres organismes.

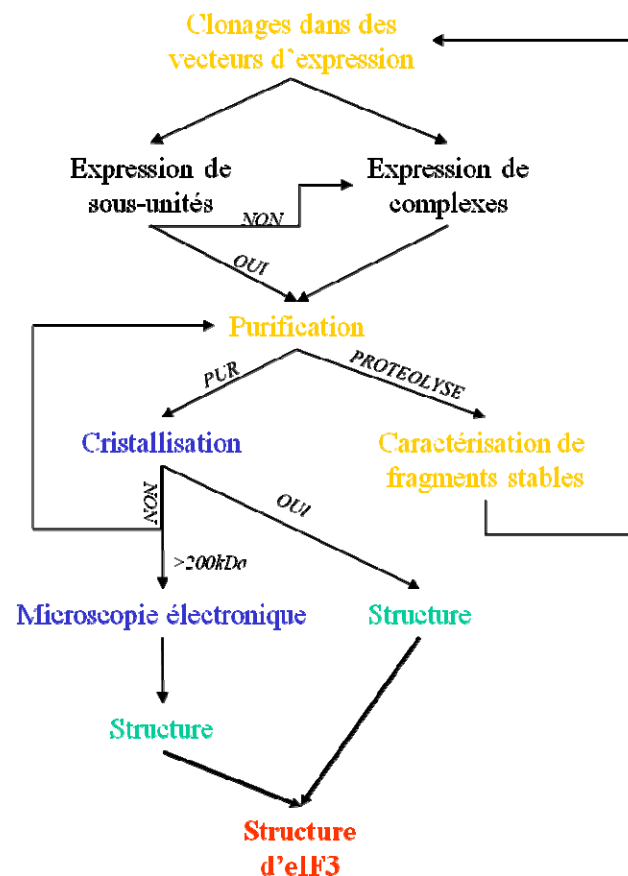


Figure 3.1 : Schéma de la démarche suivie pour la résolution de la structure du facteur de démarrage eIF3.

L'ensemble des résultats publiés sont obtenus à partir d'eIF3 purifié d'un extrait brut de l'organisme étudié. Dans ce travail, nous avons décidé d'exprimer toutes les sous-unités d'eIF3 dans un autre organisme, *Escherichia coli*, technique largement utilisée pour des productions hétérologues avec un bon rendement et pouvant permettre d'éviter les contaminations par des protéines interagissant avec la protéine d'intérêt. De plus, cette stratégie permet d'exprimer les sous-unités indépendamment les unes des autres, ce qui peut se révéler très intéressant dans l'objectif de la résolution de la structure du complexe. En effet, pouvoir ajouter ou enlever un partenaire du complexe peut, par exemple, permettre de visualiser sa position par microscopie électronique.

Dans un premier temps, nous avons testé la production des sous-unités isolées. Puis, peu à peu, nous avons mis au point des stratégies permettant l'expression et la purification de complexes de plus en plus proches d'eIF3 complet. Cette première partie vise à expliciter la stratégie, illustrée dans la **Figure 3.1**, que nous avons mise en place pour obtenir le système d'expression et de purification actuel. Les détails concernant les purifications de complexes et de sous-complexe seront abordés dans les parties suivantes.

I Clonage des gènes codant pour les sous-unités du facteur eIF3 de levure

Le génome de la levure *Saccharomyces cerevisiae* a été séquencé en avril 1996. Il s'agissait du premier génome eucaryotique séquencé en entier. Résultat de l'effort conjoint de 96 laboratoires et plus de 600 scientifiques regroupés au sein d'un consortium international, la connaissance de ce génome a notamment permis de débiter les comparaisons entre les génomes d'organismes issus des trois domaines du vivant et de mener les premières analyses sur les « gènes de cœur » universellement conservés ([Clayton et al., 1997](#)). L'identification et le clonage des ADNc des gènes humains codant pour les sous-unités d'eIF3 ont permis par alignement de ces séquences avec le génome de levure de clairement identifier cinq gènes homologues ([Danaie et al., 1995](#); [Naranda et al., 1994](#)). A mon arrivée au laboratoire, des vecteurs d'expression procaryotiques portant chacun des cinq gènes avaient été construits.

Les vecteurs pET15b (Novagen) utilisés permettent une fusion N-terminale de la protéine d'intérêt avec une étiquette hexahistidine (His₆) (**Figure 3.2**). Cet étiquetage facilite la purification par l'accrochage spécifique de la protéine sur une résine appropriée. pET15b possède une origine de réplication ColE1. Un site de clonage multiple contenant des sites uniques d'enzymes de restriction est placé en aval d'une séquence promotrice reconnue par l'ARN polymérase du bactériophage T7. Un site terminateur reconnu par cette même polymérase est présent en aval du site de clonage. Le plasmide comporte également l'opérateur lac, site de liaison du répresseur de la transcription de l'operon lactose (LacI). En absence de lactose, la liaison de LacI sur ce site contribue à réprimer la transcription du gène situé en aval. Le plasmide confère à la bactérie une résistance à l'antibiotique Ampicilline (Amp). Nous avons construit au laboratoire une version modifiée du vecteur pET15b nommée pET15blpa comportant une succession de sites de restrictions uniques additionnels au niveau du site de clonage. Ainsi, il a été ajouté entre les sites *NdeI* et *BamHI* du vecteur commercial une séquence correspondant aux sites : *SmaI/XhoI/KpnI/SacII/NotI/SalI* de telle façon à ce que le site *BamHI* soit détruit après le clonage de ces sites (**Figure 3.2**). Le clonage de la sous-unité eIF3a a été réalisé dans une autre version modifiée du vecteur pET15b notée pET15bXSB correspondant à un site de clonage comprenant les sites : *XhoI/SmaI/BamHI*.

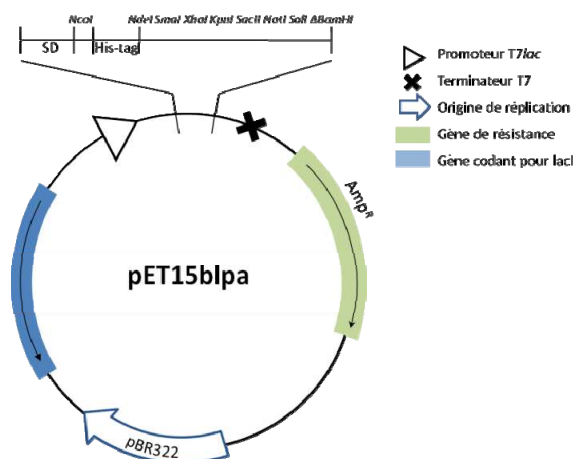


Figure 3.2 : Le plasmide pET15blpa

II Essais de surproduction

Dans un premier temps, l'expression des sous-unités isolées a été estimée par des essais de surproduction sur des cultures en petits volumes. L'usage et la fréquence des codons est variable d'un organisme à l'autre. La souche d'E.coli Rosetta™(DE3) est conçue spécialement pour la production de protéines eucaryotiques dans la bactérie (génotype : $F^- ompT hsdS_B (r_B^- m_B^-) gal dcm \lambda(DE3)$). La souche Rosetta™(DE3)-pLacIRARE contient en plus un plasmide pLacIRARE. Ce plasmide comporte le gène de résistance à l'antibiotique chloramphénicol (Cm), une origine de réplication p15a compatible avec ColE1 et oriC et le gène codant pour le répresseur LacI. D'autre part, le plasmide porte la cassette RARE contenant les gènes codant pour les ARNt portant les anticodons complémentaires des codons rares (AGG, AGA, AUA, CUA, CCC et GGA) ([Novy et al., 2001](#)). Ainsi, les cellules Rosetta™(DE3)pLacIRARE transformées par des vecteurs de type pET15b peuvent être sélectionnées par les antibiotiques Cm et Amp. Une première phase de culture des bactéries en milieu riche en présence d'antibiotique permet d'augmenter la biomasse de bactéries répliquant le plasmide. Afin d'exprimer la protéine, un analogue non-hydrolysable du lactose, l'IPTG (Isopropyl β -D-1-thiogalactopyranoside), est ajouté à la culture. Il se fixe sur la protéine LacI, levant ainsi la répression de l'opéron lactose. Cela active l'expression du gène de l'ARN polymérase du bactériophage T7, porté sur le chromosome bactérien et libère de surcroît le promoteur T7 de pET15b, occupé par LacI. L'ARN polymérase du phage T7 va alors transcrire les gènes clonés dans le vecteur pET15b. Les ARNm ainsi produits sont ensuite pris en charge par la machinerie traductionnelle de la bactérie hôte et leur abondance permet d'obtenir généralement des protéines en grande quantité.

Pour mener ces essais de purifications, 50mL de milieu riche sélectif sont inoculés puis induits. La présence d'une étiquette His₆ permet une purification en une étape afin d'estimer les niveaux de surproduction des différentes sous-unités. La résine TALON™ de type IMAC (Immobilized Metal Affinity Chromatography) utilisée pour accrocher l'étiquette est une résine basée sur des ions cobalt fixés fermement dans une poche électronégative formée par un chélatant tétradantate lié covalamment à des billes de polymère inerte. Les doublets électroniques des histidines de l'étiquette sous forme déprotonnée sont coordonnés par les lacunes électroniques de l'ion cobalt ; l'étiquette est ainsi retenue sur la résine. Le déplacement peut alors être assuré par une augmentation du pH induisant la protonation des histidines ou par l'ajout d'imidazole correspondant à la chaîne latérale de l'histidine. L'équilibre de la fixation des histidines peut donc être déplacé par compétition. Pour de tels essais, une méthode en suspension (« batch ») est utilisée, elle est plus rapide mais le produit obtenu est en général moins pur qu'en utilisant un système chromatographique classique.

L'analyse des éluions après accrochage des extraits bruts de cultures de sous-unités seules par électrophorèse en conditions dénaturantes (SDS-PAGE : Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis) montre clairement un bon niveau de surproduction des protéines eIF3i et eIF3g bien que cette dernière apparaisse sous la forme de deux bandes (**Figure 3.3**). En revanche, les éluions des cultures correspondant aux trois grosses sous-unités (eIF3a, b, c) ne font apparaître aucune protéine majoritaire et aucune surproduction n'est observée dans les surnageants de sonication. Il peut-être éventuellement constaté une surproduction dans les culots de sonication qui pourrait correspondre aux protéines dénaturées dans les corps d'inclusion (**Figure 3.3**).

Afin de vérifier si l'absence de protéines dans le cas d'eIF3a, eIF3b et eIF3c n'est pas liée à un problème de stabilité dans les extraits bruts en raison de l'absence de ses partenaires, des culots de bactéries surexprimant chacune de ces protéines ont été mélangés puis les cellules ont été lysées par sonication. Dans ce cas non plus, il n'a pas été observé de fixation spécifique. Finalement, pour éviter d'être sensible à une éventuelle protéolyse N-terminale des sous-unités, nous avons également tenté de pêcher les différents partenaires du complexe en utilisant la sous-unité 3i comme appât. Malheureusement, la présence de ces protéines n'a encore pas pu être détectée.

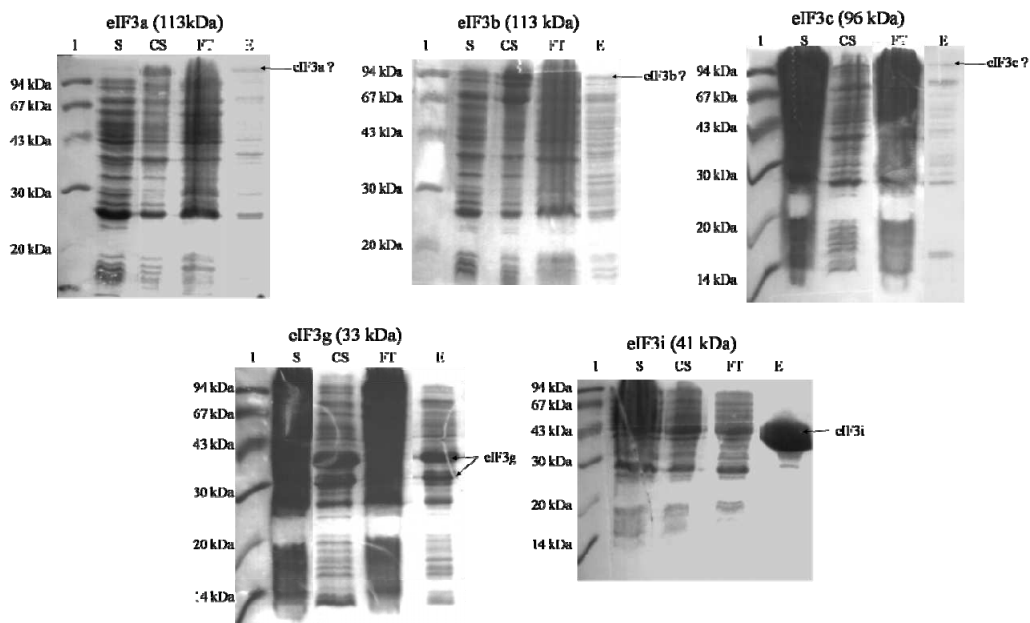


Figure 3.3 : Profils d'élution de la résine TALON des sous-unités isolées. Les échantillons correspondant au surnageant de sonication (S), au culot de sonication (CS), au surnageant de sonication qui passe au travers de la résine (FT) et à l'élution (E) ont été appliqués sur gels de polyacrylamide (12% : eIF3a, 3b, 3c ; 16% : eIF3g, eIF3i). Piste 1 : marqueur de poids moléculaire LMW (GE Healthcare).

Le dernier essai préliminaire a consisté à mélanger des culots de bactéries correspondant à 1/3 de litre de culture de chacune des cinq sous-unités. La purification par accrochage sur résine TALON a été réalisée sur un système chromatographique. L'élution a été analysée sur un gel d'électrophorèse en conditions dénaturantes. Il faut noter que la purification a été menée très rapidement (moins de 6 heures) de façon à limiter la protéolyse. Plusieurs bandes de protéines de masses comprises entre 30 et 100kDa ont été mises en évidence. Afin de confirmer la présence d'un complexe, l'élution de la résine a été appliquée sur un tamis moléculaire analytique dédié aux hauts poids moléculaires (Superdex200HR 10/30, GE Healthcare). Le profil d'élution de la colonne présente plusieurs pics et épaulements dans la région des plus hauts poids moléculaires, le poids apparent de la première espèce éluée pouvant correspondre au complexe pentamérique (environ 360kDa). L'analyse de ces différents pics par SDS-PAGE indique la présence de nombreuses bandes concentrées majoritairement dans deux régions (**Figure 3.4**). La première comprend des poids de 30 à 43kDa, zones où doivent migrer les protéines eIF3i et eIF3g, et la seconde s'étend de 76 à 116kDa et peuvent correspondre aux migrations des protéines eIF3a, eIF3b et eIF3c. Les autres pics peuvent correspondre à des sous-unités protéolysées. De plus, deux bandes migrent à la même hauteur que celles observées lors de la purification de la sous-unité 3g. Afin de mieux caractériser ce complexe, les différentes bandes ont été excisées du gel d'électrophorèse et analysées par spectrométrie de masse.

Sommairement, les bandes excisées sont soumises à une digestion à la trypsine puis les peptides produits sont extraits du gel avant d'être analysés en spectrométrie de masse. Les masses des peptides identifiés sont alors comparées à une base de données générant les peptides obtenus après l'hydrolyse trypsique de protéines issues d'organismes sélectionnés. Il convient néanmoins de préciser que les peptides utilisés dans la recherche correspondent à ceux effectivement observés sur le spectre. Les peptides les mieux extraits seront les peptides hydrophobes alors que certains peptides générés par la digestion restent dans le gel en raison de leur hydrophilie. Cette méthode permet d'identifier l'origine de la protéine étudiée mais pas de déterminer précisément les sites de protéolyse. Quelques renseignements peuvent malgré tout être obtenus en analysant la couverture de séquence offerte par les différents peptides.

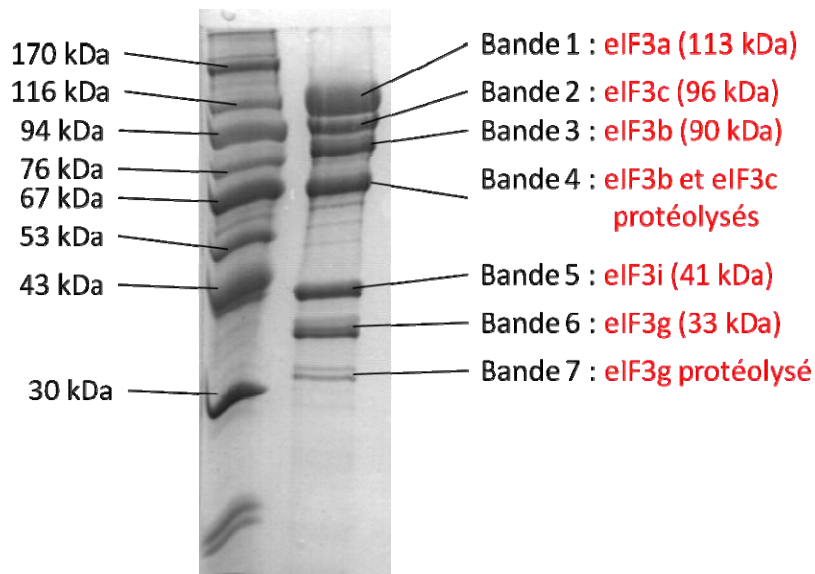


Figure 3.4 : Profil de la purification réalisée par co-cassage des 5 sous-unités. L'échantillon obtenu après passage sur Superdex 200 a été déposé sur un gel de polyacrylamide 10% en conditions dénaturantes (Piste 2). Le gel est coloré au bleu de coomassie après migration. La piste 1 correspond à un mélange des marqueurs de poids moléculaires HMW et LMW (GE Healthcare). Les bandes analysées en spectrométrie de masse sont numérotées. Les protéines identifiées sont indiquées au regard de chaque bande et leurs masses théoriques reportées entre parenthèses.

Les analyses des digestions tryptiques ont été réalisées par Anne Pascale Bouin et Guillaume Van der Rest au Laboratoire de Chimie des Milieux Réactionnels (DCMR, Ecole Polytechnique). Les masses des peptides obtenues ont été comparées avec le moteur de recherche MASCOT. De façon très satisfaisante, toutes les sous-unités ont pu être identifiées. La numérotation des résidus utilisée ci-dessous fait référence aux séquences des protéines natives.

Bande 1 : les peptides couvrent 30% de la séquence d'eIF3a, ce qui est acceptable pour une protéine de cette taille. De plus, les peptides N et C-terminaux correspondent quasiment aux extrémités de la séquence de la protéine indiquant une protéine *a priori* intacte.

Bande 2 : la seconde série de peptides couvre 37% de la séquence d'eIF3c. Le premier peptide N-terminal débute sur le résidu 130 alors que la séquence C-terminale est mieux déterminée.

Bande 3 : les peptides couvrent 25% de la séquence d'eIF3b. Le peptide N-terminal s'étend des résidus 110 à 125 et le dernier peptide est situé 50 résidus en amont de l'extrémité C-terminale.

Bande 4 : les peptides extraits de cette bande correspondent à un mélange des protéines eIF3b et eIF3c. En ce qui concerne eIF3b, bien que la séquence soit mieux couverte (37%), le premier résidu clairement identifié est le résidu 200 indiquant une possible protéolyse N-terminale. Le peptide C-terminal identifie la protéine jusqu'au résidu 720 (40 résidus avant l'extrémité C-terminale). Les peptides correspondant à la digestion d'eIF3c couvrent 26% de la séquence. Ils se concentrent dans la partie C-terminale à partir du résidu 305 mais ne définissent pas les 45 résidus C-terminaux.

Bande 5 : La cinquième bande identifie clairement la séquence d'eIF3i avec une couverture de 60%. Les peptides N et C-terminaux de la protéine sont présents.

Bande 6 : Les peptides extraits de cette bande correspondent clairement à la protéine eIF3g dont 69% de la séquence est couverte, les peptides étant répartis des extrémités N à C-terminale.

Bande 7 : Les peptides de cette bande et ceux de la bande 6 conduisent à des alignements très similaires. Ils recouvrent environ 43% de la séquence d'eIF3g débutant sur le 75ème résidu et couvrent presque parfaitement la partie C-terminale de la protéine.

Précédemment, un sous complexe comprenant les sous-unités eIF3a, eIF3b et eIF3c n'a pu être purifié. L'absence d'expression pourrait être liée à une instabilité du complexe en raison de l'absence de ses partenaires. Cependant, la littérature fait référence à ce complexe comme le cœur fonctionnel minimal du facteur eIF3 pouvant être purifié indépendamment des sous-unités eIF3i et eIF3g ([Phan et al., 2001](#)).

Une autre possibilité serait l'inaccessibilité des étiquettes polyhistidines au sein du complexe eIF3a-3b-3c, en raison d'un enfouissement de l'étiquette, ou à une dégradation N-terminale des protéines. Les purifications réalisées par la suite apportent des éléments de réponse. Il apparaît en effet qu'au sein de ce trimère, les étiquettes d'eIF3a et d'eIF3c sont enfouies. De plus, comme nous le verrons après, la fusion N-terminale ne semble pas être adaptée pour utiliser la sous-unité eIF3b comme appât.

Ces premiers résultats sont encourageants et démontrent qu'il est envisageable de purifier les cinq sous-unités du facteur de levure chez la bactérie. Néanmoins, les rendements sont très faibles puisqu'à partir d'1,5 litre de culture et après seulement deux étapes de purification, il est obtenu environ 0,5 mg de complexe. De plus, il apparaît clairement une protéolyse de certaines des sous-unités. Dans un premier temps, nous avons donc cherché à isoler des fragments stables de chacune des sous-unités afin d'améliorer la pureté et le rendement.

III Mise en place d'une stratégie de coexpression

Afin de pallier les problèmes exposés ci-dessous, les stratégies de coexpression permettant de surproduire plusieurs protéines partenaires au sein d'une même cellule sont de plus en plus utilisées. En effet, cette démarche peut en certains cas améliorer considérablement les rendements d'expression d'une protéine si celle-ci est produite en présence de son partenaire.

Le principe sur lequel repose la technique est simple. Une bactérie est transformée par plusieurs plasmides exprimant des gènes de résistance à des antibiotiques différents, comportant des origines de réplication compatibles et, bien entendu, les différents gènes de nos protéines. La présence de gènes de résistance différents permet une sélection efficace des transformants alors que des origines de réplication compatibles assurent une réplication indépendante des plasmides les uns par rapport aux autres. Des résultats récents montrent néanmoins que des plasmides portant des origines de réplication identiques peuvent malgré tout cohabiter et se répliquer pendant un nombre de générations suffisant pour assurer l'expression des protéines correspondantes ([Velappan et al., 2007](#)).

Notre but étant de coexprimer cinq protéines, nous nous sommes tournés vers le système commercial de Duet Vectors proposé par Novagen. Le système Duet propose de coexprimer deux protéines à partir d'un même plasmide. Pour cela, deux groupes de sites de restriction (Multi-Cloning Sites : MCS1 et MCS2) permettent de cloner deux gènes placés en tandem. Chaque gène est placé sous contrôle d'un promoteur T7lac correspondant à la combinaison d'un promoteur pour l'ARN polymérase du bactériophage T7 et du site opérateur de l'opéron lactose. Le plasmide contient un unique site de terminaison pour l'ARN polymérase du bactériophage T7. Le vecteur possède de plus une séquence codant pour une étiquette His₆ en amont du premier site de clonage. Le deuxième site de clonage comporte quant à lui une séquence codant pour une autre étiquette en C-terminal. Nous n'avons pas jugé utile d'utiliser cette étiquette C-terminale. Le jeu de 4 vecteurs proposé par Novagen propose des variations sur l'origine de réplication et sur les résistances aux antibiotiques. Dans notre cas, nous prévoyons de coexprimer au maximum 5 protéines, il nous est donc nécessaire de combiner au moins deux vecteurs de type Duet et un des vecteurs de type pET dont nous disposons déjà. Ainsi, le choix des vecteurs Duet s'est porté sur les vecteurs à origine de réplication compatibles, pCDFDuetTM-1 et pCOLADuetTM-1. Les vecteurs pCDFDuetTM-1 et pCOLADuetTM-1 comportent les origines de réplifications CloDF13 et ColA respectivement, compatibles avec les origines oriC, p15a et ColE1 et les gènes de résistances aux antibiotiques Spectinomycine (Sm) et Kanamycine (Kan) respectivement. A la place du vecteur pCOLADuetTM-1, le vecteur pRSFDuetTM-1 basé sur une origine de réplication

RSF1030 aurait pu être utilisé. En outre, l'origine de réplication permet à ce vecteur d'être répliqué en plus grand nombre (plus de 100 copies par cellule à comparer avec les 20 à 40 copies assurées par l'origine ColA). Il a néanmoins été observé qu'en combinaison avec des vecteurs de type pET et pCDF, les niveaux d'expression de tous les partenaires sont plus élevés avec le vecteur pCOLADuetTM-1 comparés à ceux obtenus en combinaison avec pRSFDuetTM-1, déterminant ainsi notre choix (Held et al., 2004).

Le choix des clonages à réaliser, en termes de placement des gènes et troncatures, résulte de la combinaison des résultats préliminaires que nous avons obtenus avec les connaissances bibliographiques.

En premier lieu, le schéma d'interaction déduit des résultats biochimiques propose un rôle central à eIF3b qui lie les quatre autres (eIF3a, 3c, 3i et 3g). Cependant, eIF3i et 3g interagissent entre elles mais pas avec 3a ou 3c (**Figure 3.5a**). Ainsi, le choix de cloner eIF3b dans le vecteur de type pET permettra d'envisager facilement la purification de sous-complexe à 3 partenaires. De plus, les interactions du facteur eIF3i et eIF3g au sein du complexe incitent à les placer en tandem sur un plasmide Duet en tant que partenaires plus « intimes ». De la même façon, les sous-unités eIF3a et eIF3c ont été clonées dans un même vecteur Duet. Les trois vecteurs étant répliqués à des nombres de copies similaires (environ 40 copies en théorie), un choix arbitraire a conduit à opter pour le vecteur pCOLADuetTM-1 pour exprimer eIF3i et eIF3g alors que les ORF codant pour les sous-unités eIF3a et eIF3c ont été clonées dans le vecteur pCDFDuetTM-1.

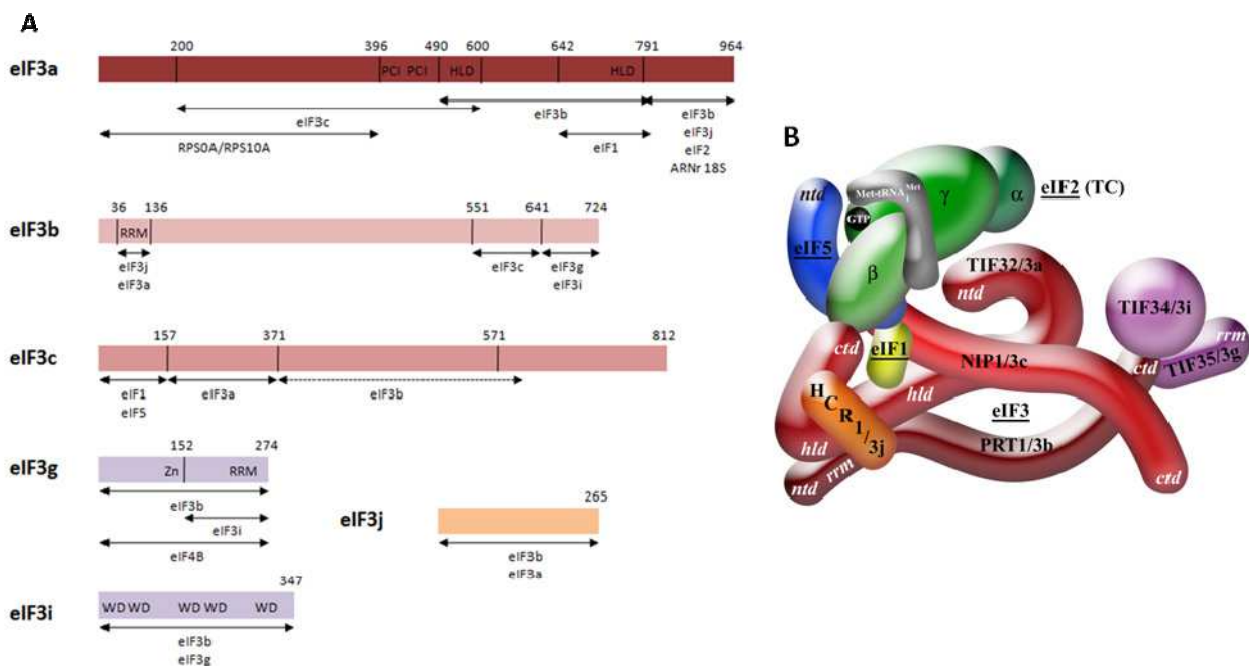


Figure 3.5a : Organisation du MFC par le facteur de démarrage eIF3 d'après (Valasek et al., 2002). (A) Représentation schématique des interactions répertoriées pour les différentes sous-unités du facteur eIF3. Les interactions d'eIF3a avec l'ARNr et les protéines ribosomales S0A et S10A sont indiquées. (B) Vue schématique du complexe multifacteurs.

Le vecteur pET15bpa-eIF3b permet de produire la sous-unité eIF3b avec une étiquette His₆ en N-terminal. La position de l'étiquette peut avoir deux effets majeurs : elle peut empêcher la formation d'interactions proches de l'étiquette et, si une protéolyse se produit par cette extrémité, la purification par IMAC devient inopérante. Ainsi, après les résultats médiocres obtenus pour la surproduction de la sous-unité seule et ceux obtenus en SM laissant supposer une protéolyse N-terminale, il a été décidé de sous-cloner le gène codant pour eIF3b dans un vecteur pET32b permettant l'expression d'une protéine en fusion C-terminale avec une étiquette His₆.

En outre, l'analyse en spectrométrie de masse indique une couverture de la séquence des deux bandes correspondant à eIF3b quasiment complète pour l'extrémité C-terminale alors que, du côté N-terminal, les premiers résidus identifiés laisseraient penser à une délétion. Celle-ci semble de plus correspondre à la troncature du motif RRM (résidus 36 à 136) impliqué dans la liaison d'eIF3a et eIF3j. Finalement, d'après la littérature, un sous-complexe stable peut être formé entre l'extrémité C-terminale d'eIF3b et les protéines eIF3i et eIF3g. Ainsi, il serait également intéressant de cloner cette autre sous-région d'eIF3b.

L'ensemble de ces éléments conduit à envisager plusieurs formes tronquées de la protéine (Figure 3.6) :

3bN1 : correspondant aux résidus 1 à 175 de la protéine, liant uniquement eIF3a

3bC1 : troncature de la partie N-terminale jusqu'au résidu 193. Cette forme tronquée contient les sites d'interaction avec eIF3c (551 à 641) et eIF3g et eIF3i (83 résidus C-terminaux). Elle peut lier eIF3a via l'interaction eIF3a:eIF3c

3bC2 : partie C-terminale (du résidu 589 au résidu 724) permettant la formation d'un sous-complexe de taille minimale eIF3bC2:eIF3g:eIF3i. Quelques résidus additionnels C-terminaux d'eIF3b ont été conservés en amont du site d'interaction avec eIF3g et eIF3i afin de conférer une meilleure stabilité au fragment et notamment par inspection du repliement prédit par la répartition des résidus hydrophobes (Gaboriaud et al., 1987).

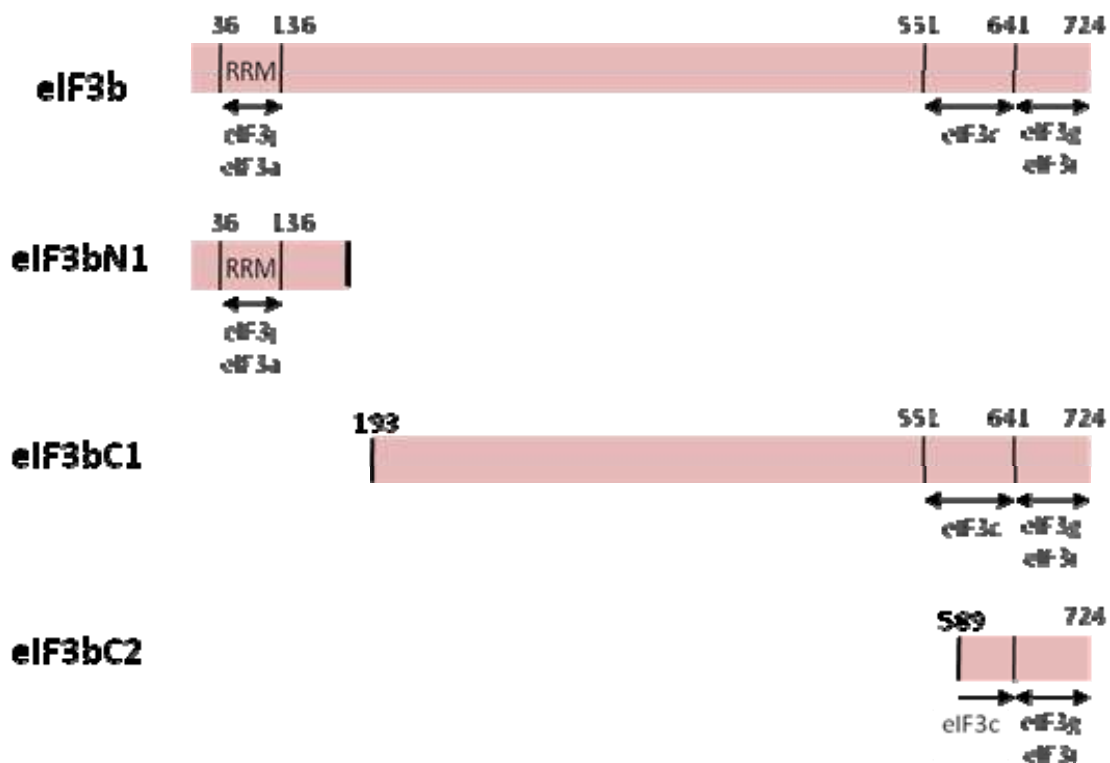


Figure 3.5b : Représentation schématique de la sous-unité eIF3b et des différentes troncatures envisagées.

Une bande correspondant à une protéolyse d'eIF3c (bande 4) a également été identifiée. Clairement, la protéolyse doit avoir lieu à l'extrémité N-terminale puisque la séquence n'est couverte pour les deux bandes d'eIF3c qu'à partir des résidus 130 et 305. Or, par l'intermédiaire de l'extrémité N-terminale d'eIF3c (résidus 1-157) se lient les partenaires du MFC (eIF1, eIF2 et eIF5). L'absence de cette portion pourrait ainsi être liée à l'absence de ces interactions stabilisantes.

Ainsi, des troncatures du facteur ont été envisagées (**Figure 3.6**) :

3cN1 : portion N-terminale (résidus 1 à 206) permettant de former un MFC sans aucune autre sous-unité d'eIF3 : eIF3cN1:eIF1:eIF2:eIF5.

3cC1 : troncature de la partie N-terminale avant le résidu 158 conservant les sites de liaison avec eIF3a (157 à 371) et eIF3b (371 à 570) mais pas les sites de formation du MFC.

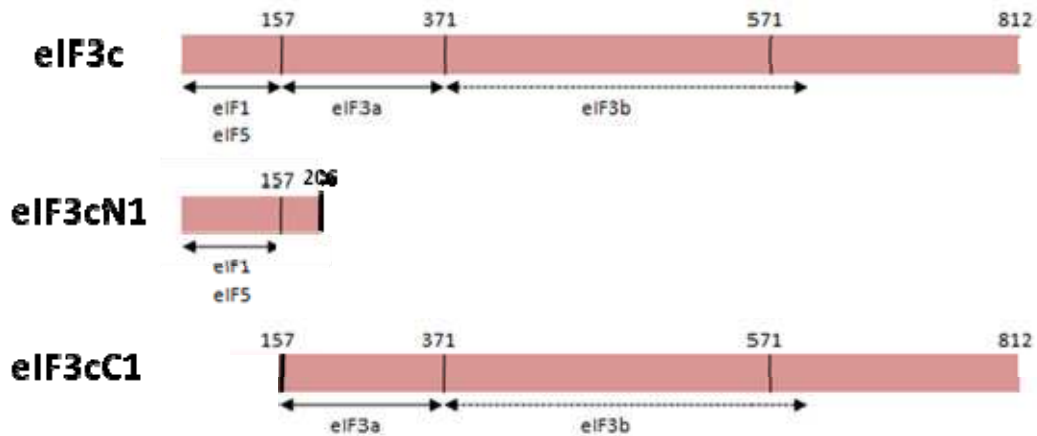


Figure 3.6 : Représentation schématique de la sous-unité eIF3c et des différentes troncatures envisagées.

La sous-unité eIF3i apparaît clairement intacte, elle ne sera donc pas modifiée. La sous-unité eIF3a ne semble pas dégradée et apparaît sous forme d'une bande unique n'indiquant à nouveau pas de troncature intéressante. Enfin, la sous-unité eIF3g est apparue au cours des différents essais sous la forme de deux bandes. La caractérisation de ces deux bandes sera décrite plus loin.

IV Clonages dans les vecteurs Duet et résultats préliminaires de la coexpression

1 Clonages

Le choix des protéines qui porteront une étiquette His6 au sein des vecteurs Duet a largement été conditionné par les sites de restrictions présents dans les ORF codant pour les sous-unités. En effet, de nombreux sites de clonages de ces vecteurs étant présents à l'intérieur des gènes codant notamment pour les sous-unités eIF3a et eIF3b. Lorsque des troncatures ont été envisagées, les formes natives des protéines ont également été clonées (**Figure 3.7**).

pET32b

Le clonage des ORF d'eIF3b dans le vecteur pET32b a été réalisé par amplification des fragments d'intérêt à partir du vecteur pET15blpa-eIF3b à l'aide d'oligonucléotides comportant les sites *NdeI* en amont du gène et *XhoI* en aval.

pCDFDuet™-1

Le gène codant pour la sous-unité eIF3a a été cloné dans le MCS1 de pCDFDuet™-1 entre les sites *BamHI* et *NotI*. Cette sous-unité sera donc produite fusionnée à une étiquette His₆ en N-terminal.

Les gènes correspondant aux différentes formes d'eIF3c ont été insérées entre les sites *KpnI* et *XhoI* du MCS2 de pCDFDuet™-1. Le plasmide pCDF-3a-3cC1 a été construit à partir de pCDF-3a-3c par excision dirigée par des oligonucléotides. Notons que la même technique a été utilisée pour construire pET15blpa-3cC1 à partir de pET15blpa-eIF3c.

pCOLADuet™-1

La sous-unité eIF3i a été choisie pour recevoir l'étiquette N-terminale, son gène a donc été cloné dans le MCS1 entre les sites *BamHI* et *NotI*.

Le clonage des gènes codant pour les différentes formes d'eIF3g a été réalisé à l'aide des enzymes de restriction *MfeI* et *XhoI*.

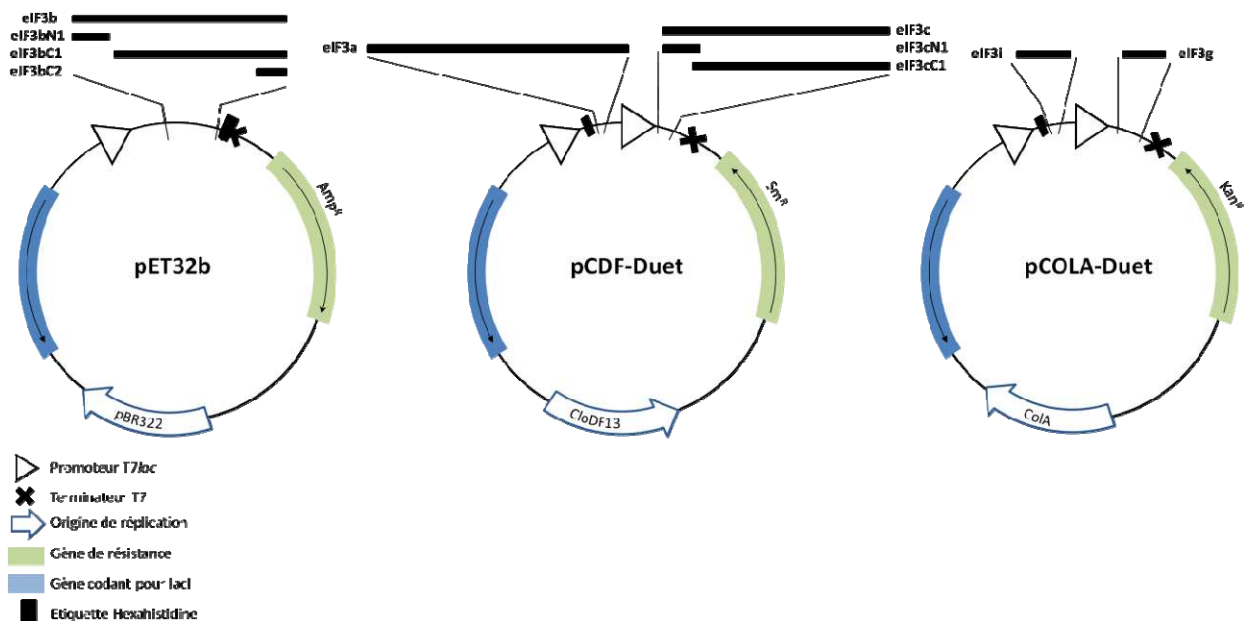


Figure 3.7 : Représentation schématique des clonages réalisés dans le système Duet.

2 Purification du complexe eIF3 dans le système de coexpression

Nous avons alors voulu voir si la coexpression pouvait permettre de limiter les protéolyses constatées lors des essais en cocassage. Les vecteurs pET32b-3b, PCDF-3a-3c et pCOLA-3i-3g ont ainsi été co-transformés dans la souche Rosetta™(DE3)-pLacI RARE. La purification a été réalisée en trois étapes (IMAC, échange d'ions et tamis moléculaire). Un complexe de haut poids moléculaire a pu être isolé sur un tamis moléculaire bien que la stœchiométrie apparaisse déficiente en la sous-unité eIF3a (**Figure 3.8**). La protéolyse d'eIF3g est toujours observée. De plus, dans la région des protéines eIF3b et eIF3c il n'apparaît qu'une bande unique qui migre à la même hauteur que la bande précédemment identifiée comme le

mélange des produits de protéolyse de ces deux protéines. En outre, des bandes sont apparues sous la bande majoritaire précédemment identifiée comme eIF3a. Les bandes pressenties comme appartenant à eIF3a et la bande correspondant au mélange des sous-unités eIF3b et eIF3c protéolysées ont été caractérisées en SM et en séquençage N-terminal. Le séquençage N-terminal montre sans ambiguïté que les deux bandes situées sous eIF3a sont celles de protéines qui débutent par la séquence GSSHHH correspondant au début d'une étiquette histidine. Celles-ci ne peuvent provenir que d'eIF3a puisque, dans cette zone de haut poids moléculaires, seules eIF3b et eIF3c pourraient être identifiées. Or, eIF3b comporte une étiquette His₆ à son extrémité C-terminale et eIF3c est dénué d'étiquette. Ainsi, il apparaît qu'eIF3a subisse une protéolyse par son extrémité C-terminale. Les troncatures envisagées seront alors un compromis entre les peptides C-terminaux observés en SM et le schéma d'interaction de la littérature.

Cette purification fait apparaître une forte protéolyse d'eIF3a alors que les sous-unités eIF3b et eIF3c intactes ne sont pas détectées. Mis à part le système d'expression, la seule différence entre cette purification et celle réalisée en co-cassage réside dans le temps de purification. En effet, la première purification a été réalisée en moins de 6 heures, en deux étapes chromatographiques. Le pic de haut poids moléculaire du tamis moléculaire a de plus été appliqué sur gel SDS-PAGE peu après son élution. En revanche, les quantités de complexe obtenues en coexpression nous ont fait envisager une étape supplémentaire de purification et l'application sur un tamis moléculaire préparatif (Superdex 200 16/60, GE Healthcare). Ainsi, la purification a duré deux jours et l'analyse sur gel SDS-PAGE du complexe purifié du pic de haut poids moléculaire révèle une protéolyse plus importante, comparée à l'analyse du pic élué de la première étape IMAC (**Figure 3.8(A) et (B)**).

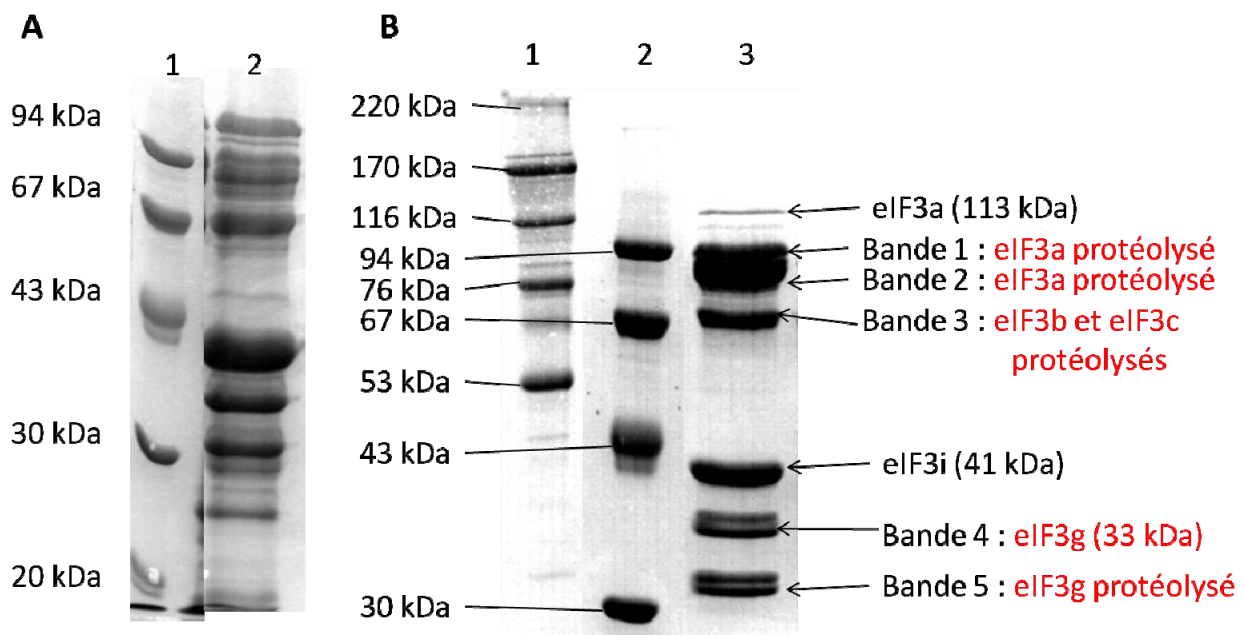


Figure 3.8 : Profil de la purification par coexpression des 5 sous-unités. (A) Les échantillons obtenus après l'étape IMAC (A, Piste 2) ou après Superdex 200 (B, Piste 3) ont été déposés sur des gels de polyacrylamide 10% en conditions dénaturantes (Piste 2). Les gels sont colorés au bleu de coomassie après migration. (A) Piste 1 : marqueur de poids moléculaires LMW (GE Healthcare). (B) Piste 1 : marqueur de poids moléculaires HMW (GE Healthcare) ; Piste 2 : marqueur de poids moléculaires LMW (GE Healthcare). Les bandes analysées en spectrométrie de masse et en séquençage N-terminal sont numérotées. Les protéines identifiées sont indiquées en rouge.

3 Caractérisation d'eIF3g

La présence dans toutes les purifications d'eIF3g sous la forme de deux bandes (**Figure 3.8, bandes 4 et 5**) nous a incités à caractériser un peu plus la séquence des bandes. L'analyse par spectrométrie de masse était compatible avec une protéolyse par l'extrémité N-terminale. Pour cela, nous avons combiné une nouvelle analyse en SM et une analyse de séquence N-terminale réalisé à la plateforme d'Analyse et de Microséquençage des Protéines de l'Institut Pasteur à Paris. Deux gels SDS-PAGE ont été préparés en parallèle en y ayant déposé des échantillons présentant les deux bandes d'eIF3g. Sur l'un des gels les deux bandes ont été excisées puis soumises au traitement requis pour l'analyse en SM et l'autre gel a été transféré sur une membrane de type PVDF afin de pouvoir être analysé en séquençage N-terminal. L'analyse des spectres de SM obtenus révèle une meilleure couverture de l'extrémité N-terminale : les premiers peptides de la bande inférieure et de la bande supérieure débutent respectivement sur les 19ème et 42ème résidus de la séquence de la protéine produite à partir de pET15blpa-eIF3g. Les résultats du séquençage N-terminal montrent finalement que la bande inférieure correspond à une protéolyse N-terminale, faisant démarrer la protéine sur le 66ème résidu (correspondant au 47ème résidu de la séquence de la protéine native). Le premier peptide en SM de la bande inférieure pouvait donc être mal attribué, le second identifié pour la bande supérieure débute quant à lui sur le 82ème résidu.

De plus, les bandes d'eIF3g ne migrent pas sous la forme d'une bande bien résolue mais plutôt légèrement dédoublée (**Figure 3.8(B)**). La séquence révèle un domaine putatif de liaison du zinc, N-terminal par rapport au motif RRM (**Figure 3.4(B)**). Nous avons déjà constaté au laboratoire une migration aberrante sur gel SDS-PAGE de protéines liant le zinc (et notamment eIF2 β). En effet, si le zinc est fortement lié, la dénaturation au SDS pourrait être moins efficace. Dans ce cas, les deux bandes observées sont un artefact. Afin de tester l'échange possible du zinc par le cobalt de la résine TALON™, des essais de purification en une étape IMAC ont été réalisés avec la résine au cobalt précédemment utilisée, d'une part, et une résine basée sur la coordination de l'étiquette par des ions nickel : Ni-NTA His•Bind® (Novagen). A nouveau, un effet comparable est observé dans les deux éluions. Dans ce contexte, il faut remarquer que l'ajout d'EDTA avant la dénaturation des échantillons appliqués sur gel SDS-PAGE permet d'obtenir les deux bandes mieux résolues (**Figure 3.9**).

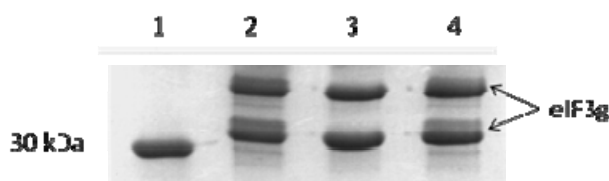


Figure 3.9 : Effet de l'ajout d'EDTA sur la migration d'eIF3g sur gel SDS-PAGE (12% acrylamide). Piste 1 : marqueur de poids moléculaires LMW (GE Healthcare). Piste 2 : ajout de 50 μ M ZnCl₂ avant dénaturation des échantillons. Piste 3 : ajout d'1mM EDTA avant dénaturation. Piste 4 : ajout de 100mM 2-mercaptoéthanol avant dénaturation. La migration des bandes correspondant à eIF3g est clairement plus homogène sur la piste 3.

Nous avons utilisé la spectrométrie d'absorption atomique pour déterminer le contenu en zinc de la protéine. Rapidement, la méthode consiste à chauffer un échantillon à une température suffisante pour assurer son atomisation par passage dans une flamme d'une température de l'ordre de 2000 à 3000K, qui fournit alors les éléments de l'échantillon à l'état fondamental. Ces éléments vont alors pouvoir absorber des photons dont l'énergie est caractéristique de l'élément et la quantité sera proportionnelle au nombre d'atomes présents dans l'échantillon. Dans notre cas, l'élution de la résine échangeuse d'ion est répartie en deux lots dont un est dialysé contre un tampon en présence d'EDTA (1mM). Comme l'indique le **tableau 3.1**, la stœchiométrie apparente est de 0,5 zinc par molécule. Il faut toutefois noter que la concentration en protéine est évaluée à partir du coefficient d'extinction molaire calculé à partir de la séquence de la protéine. Un défaut de zinc dans le milieu paraît peu probable puisqu'il est notamment remarqué que l'utilisation de verrerie induit elle-même la présence de zinc (environ 1µM) dans les solutions. De plus, la présence de ce zinc a clairement pu être mise en évidence en SM. Cette analyse fait apparaître un pic majoritaire de la protéine liée au zinc et un faible pic correspondant à la protéine seule. Les résultats ne sont cependant pas quantitatifs et le zinc peut être arraché à la protéine au moment de l'ionisation. L'interprétation la plus raisonnable est que la stœchiométrie est de 1 et que le coefficient d'extinction molaire calculé pour eIF3g est entaché d'une erreur importante.

Echantillon	Transmission	Erreur
Tampon	0,0101	0,0035
1,25µM Zn ²⁺	0,0541	0,0027
2,5µM Zn ²⁺	0,1241	0,0042
5µM Zn ²⁺	0,2404	0,0062
10µM Zn ²⁺	0,4411	0,0063
25µM Zn ²⁺	0,7996	0,0032
10µM eIF3gC1	0,2561	0,0072
35µM eIF3gC1	0,6874	0,0174
10µM eIF3gC1 + 1mM EDTA	0,2428	0,0086

Tableau 3.1 : Analyse en spectrométrie de flamme du contenu en zinc de la protéine eIF3gC1. Les valeurs de transmission de la gamme de zinc sont reportées. Le signal correspondant à une concentration de 10µM en protéine, en présence ou en l'absence d'EDTA, est identique à celui d'une solution à 5µM en zinc, indiquant une stœchiométrie ½.

V Implémentations de la stratégie de purification

1 Clonages

Différentes formes tronquées d'eIF3a (**Figure 3.10**) :

eIF3aN1 : correspond au fragment 1-918 préservant le maximum de résidus correspondant au dernier peptide détecté par les analyses en SM. Les différentes formes ont été obtenues par mutagenèse de délétion. Cette technique a été préférée à la mutagenèse ponctuelle modifiant le codon du résidu C-terminal choisi en codon STOP car elle est tout aussi efficace et rapide, tout en diminuant la taille des plasmides produisant les formes tronquées. Les plasmides pCDF-3aN1, pCDF-3aN1-3cC1 et pET15bXSB-eIF3aN1 ont ainsi été obtenus.

eIF3aN2 : s'arrête sur l'alanine 865. Cette forme a été choisie sur la base d'une prédiction de domaines stables par analyse du contenu en résidus hydrophobes ([Gaboriaud et al., 1987](#)). De la même façon que précédemment, les plasmides pCDF-3aN2, pCDF-3aN2-3cC1, pET15bXSB-3aN2 ont été construits.

eIF3gC1 : Le site de protéolyse identifié par séquençage N-terminal conduit à produire la protéine à partir du résidu leucine 47. Cette analyse a abouti à la construction des plasmides pET15bIpa-3gC1 et pCOLA-3i-3gC1 par mutagenèse de délétion.

2 Caractérisation d'eIF3bC1

La fusion en C-terminal de l'étiquette His₆ à la forme tronquée d'eIF3b a permis de purifier cette protéine isolée ou au sein de complexe. De plus, puisque la construction du vecteur pET32b ne comporte pas le site de coupure de la thrombine, nous avons été amenés à l'ajouter au plasmide pET32b-3bC1 afin de pouvoir éventuellement retirer l'étiquette.

3bC1thr : L'ajout d'un site de coupure de la thrombine (LVP↓RGS) a été réalisé par mutagenèse d'insertion. Le plasmide ainsi obtenu est noté pET32b-3bC1thr

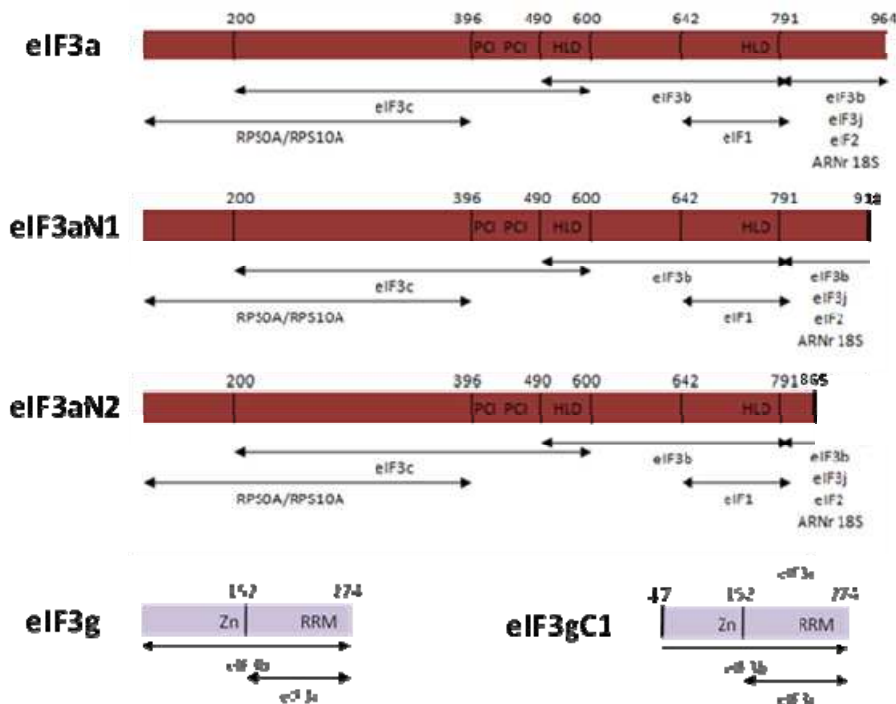


Figure 3.10 : Représentation schématique des sous-unités eIF3a et eIF3g et des différentes troncutures envisagées. Les numérotations correspondent aux séquences des protéines natives.

3 Stœchiométrie des complexes

Plusieurs observations apparues au cours des différentes purifications nous ont conduits à envisager des modifications des constructions plasmidiques. Ces remarques concernent essentiellement la stœchiométrie des différents complexes. En premier lieu, il apparaît que la sous-unité eIF3a est faiblement surproduite dans le système de coexpression alors qu'exprimée seule, les niveaux de surproduction sont très bons. Nous avons donc tenté de produire eIF3a à partir du plasmide pET15blpa-eIF3aN1 ou pET15blpa-eIF3aN2 alors qu'eIF3bC1 serait produite à partir d'une nouvelle construction en tandem avec eIF3cC1.

L'étiquetage polyhistidine était nécessaire dans la phase de démarrage du projet, afin de détecter le plus de sous-complexes possibles. Dans un deuxième temps, il est devenu préférable de n'étiqueter qu'une protéine afin d'obtenir une stœchiométrie uniforme. Nous avons donc entrepris d'enlever les étiquettes sur tous les plasmides Duet et pET utilisés dans la coexpression. La sous-unité eIF3bC1 toujours présentée comme le meilleur partenaire pour mener les expériences de « pull-down » est le choix le plus évident pour l'étiquetage. Toutefois, les sous-unités eIF3a et eIF3c étant peu produites, elles sont potentiellement intéressantes pour porter l'étiquette afin de sélectionner un complexe de bonne stœchiométrie. La bibliothèque de plasmides ne comportant pas de vecteur permettant de produire eIF3cC1 sous forme fusionnée à une étiquette, il a donc été décidé de le construire. Les sous-unités eIF3i et eIF3gC1 n'ont pas été retenues en raison de leurs plus forte surexpression au regard des autres sous-unités du complexe.

Clonage d'eIF3bC1 dans un vecteur Duet :

un clonage par enzymes de restriction entre les sites *NcoI* et *NotI* du plasmide pCDF-3cC1 a été réalisé en utilisant pour matrice de PCR le vecteur pET32b 3bC1thr

Suppression des étiquettes His₆ :

eIF3i : les étiquettes ont été supprimées par mutagenèse de délétion permettant de produire la protéine sous sa forme native. Le vecteur résultant est noté pCOLA-3i(nt)-3gC1

eIF3a : le choix de produire eIF3a depuis un vecteur de type pET a conduit à supprimer les étiquettes histidines des gènes codant pour les différentes formes tronquées de la protéine. Les plasmides obtenus sont appelés pET15bXSB-3aN1(nt) et pET15bXSB-3aN2(nt)

eIF3bC1 : un plasmide pCDF-3bC1(nt)-3cC1 permet de produire les protéines eIF3bC1 et eIF3cC1 dénuées d'étiquettes.

eIF3cC1 : l'ajout d'une étiquette His₆ a été entrepris par mutagenèse d'insertion. Le plasmide produit est noté pCDF-3bC1(nt)-3cC1(t).

VI Conclusion

Cette démarche a permis d'obtenir une bibliothèque de vecteurs, résumée dans le **tableau 4.5 (Figure 3.11)**, offrant la possibilité de produire une grande variété de complexes, sous-complexes et différentes formes des sous-unités du facteur de démarrage eIF3. Nous allons maintenant nous focaliser sur les productions de complexes et sous-complexes à l'aide de ce système de coexpression.

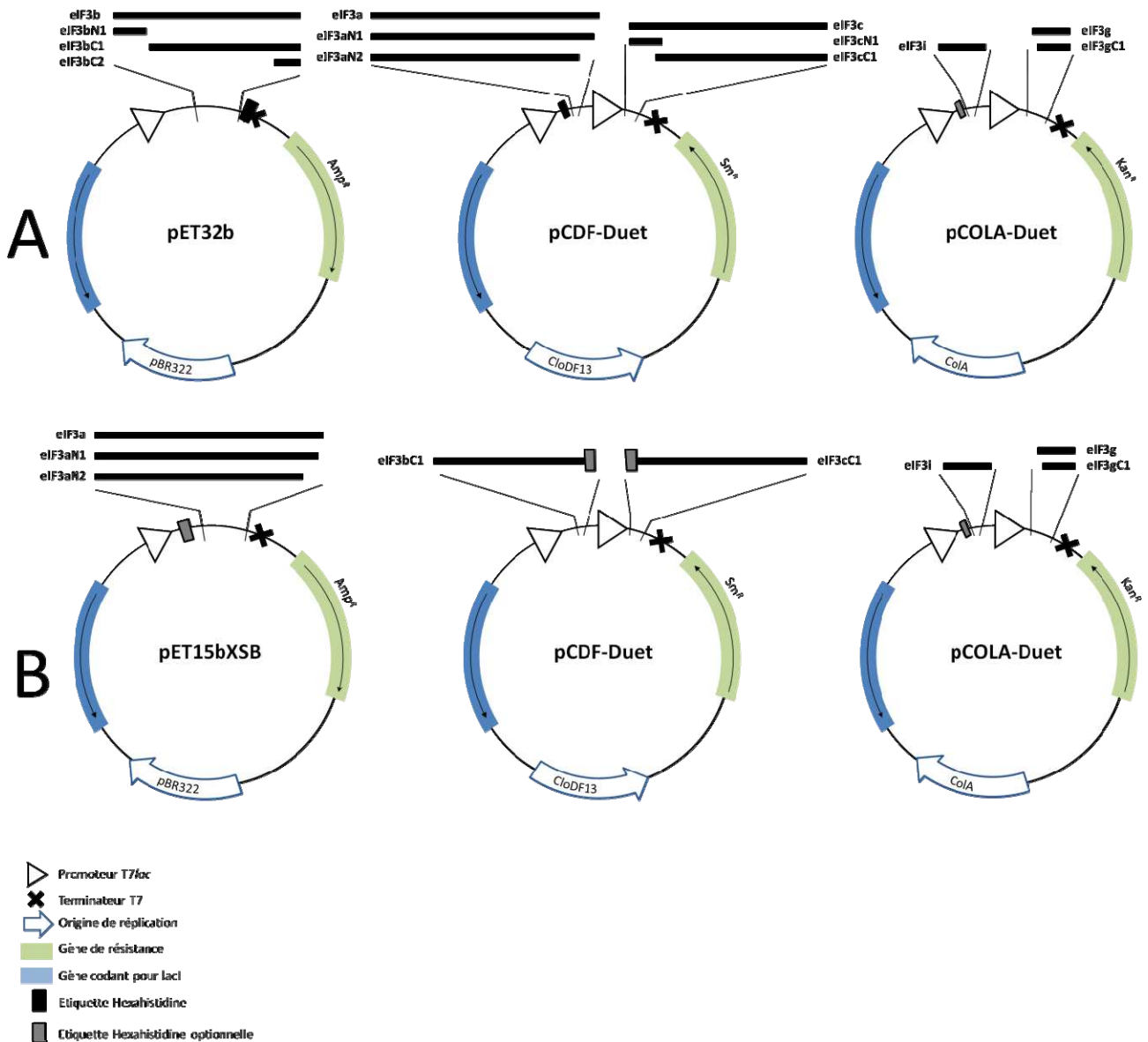


Figure 3.11 : Représentation des deux systèmes de coexpression (A et B) envisagés.

PARTIE B PURIFICATIONS DE DIFFERENTES FORMES DU FACTEUR EIF3

I Stratégie d'expression

1 Expression des protéines

La transformation d'une souche d'*E. coli* par plusieurs vecteurs d'expression nécessite quelques précautions. Tout d'abord, nous avons constaté que la co-transformation par électroporation est une méthode efficace. D'autre part, nous n'avons pas noté de problèmes de croissance bactérienne même en présence de 4 antibiotiques présents chacun à la même concentration que celle à laquelle ils sont utilisés individuellement. Afin de limiter la sélection de mutations sur les plasmides, avantageuses pour la bactérie et nuisible à l'expression des protéines d'intérêt, nous avons choisi d'ensemencer la culture en milieu liquide directement à partir des bactéries transformées resuspendues. Enfin, la température lors de l'induction est un paramètre important pour la production de protéines eucaryotiques chez *E. coli*. Il apparaît que les sous-unités sont bien produites à 20°C ; baisser la température ne semble pas apporter d'amélioration alors que l'augmenter amène à augmenter la production des protéines dans les corps d'inclusion.

Ainsi, la séquence habituelle au cours de ce travail pour la production de protéines isolées ou de complexes consiste à électroporer des bactéries Rosetta™(DE3)-pLacIRARE en présence du ou des plasmides requis. Suivant le choc électrique, les bactéries resuspendues dans un milieu non-sélectif sont laissées à 37°C pendant une heure pour les co-transformations à 3 plasmides afin de permettre l'expression des gènes de résistance portés par les différents plasmides. Le milieu liquide sélectif est ensuite inoculé puis incubé à 37°C pour une nuit. Le lendemain, la culture en fin de phase exponentielle est induite par ajout d'1mM d'IPTG pendant 4 à 5 heures à 20°C, suite à quoi la culture est centrifugée et le culot de bactéries conservé à -20°C.

2 Purification

La procédure générale de purification mise en œuvre est détaillée dans la partie Matériels et Méthodes. Sommairement, les bactéries sont resuspendues puis les parois cellulaires sont rompues par sonication dans un tampon contenant un cocktail antiprotéases, 3mM 2-mercaptoéthanol et 500mM NaCl. Une telle force ionique s'est révélée nécessaire dans la plupart des cas pour assurer la solubilité des protéines d'intérêt, et permettre de diminuer les interactions non spécifiques avec la résine TALON. Le lysat clarifié est ensuite injecté sur une colonne contenant la résine TALON. L'étiquette de la protéine éluée peut alors être éventuellement enlevée par hydrolyse par la thrombine lors d'une dialyse. Cette dialyse permet d'enlever l'imidazole, connu pour inhiber l'action de la thrombine. La protéine sans étiquette est alors extraite par un deuxième passage sur la résine TALON. L'étape suivante consiste à diminuer la force ionique de l'échantillon. Ceci a permis dans tous les cas l'accrochage des protéines sur une résine échangeuse d'anion de type Q-Sépharose. Enfin, les protéines ou complexes sont purifiés par tamisage moléculaire. Une étape supplémentaire de chromatographie d'interaction hydrophobe a été ajoutée à certaines purifications afin d'améliorer la pureté et/ou la stœchiométrie des complexes. Dans ce type de purification, les protéines précipitent sur la colonne en présence de sulfate d'ammonium. Un gradient vers un tampon sans sulfate d'ammonium permet de resolubiliser les protéines.

Toutes les purifications ayant abouti à un niveau de pureté suffisant et présentant une stœchiométrie des sous-unités présentes acceptable ont été soumises à des tests de cristallisation. Nous disposons pour cela au laboratoire d'un robot Genesis (TECAN) permettant de réaliser des gouttes de cristallisation par mélange, par exemple, d'1 μ L d'échantillon avec 1 μ L de conditions de cristallisation. Les différentes matrices de conditions de cristallisation à notre disposition permettent de tester 912 conditions différentes. La procédure de cristallisation et les principes attenants à la résolution de structure à partir de clichés de diffraction aux rayons X seront plus explicitement présentés dans la [partie D](#).

II Résultats des purifications

1 Les sous-unités isolées

Les purifications des sous-unités isolées sont résumées dans le **tableau 3.2**. Globalement, les troncatures envisagées ont permis d'obtenir pour chaque sous-unité des protéines homogènes avec de bons rendements, qui ont pu être testées pour la cristallisation. La seule sous-unité qui s'est révélée instable est eIF3a. La forte protéolyse de cette sous-unité est observée pour les deux troncatures envisagées. La majeure partie de la protéine native se trouve dans le culot de sonication et la faible fraction soluble n'est pas retenue sur la résine TALON™.

De plus, les essais de surproduction de la sous-unité eIF3bC2 seule n'ont pas permis de détecter la protéine. Cette forme tronquée de la protéine a été envisagée afin de purifier un sous-complexe minimal eIF3bC2-i-gC1. La coexpression de ces trois sous-unités n'a pas permis de détecter la formation de complexe. La protéine eIF3bC2 a néanmoins pu être purifiée mais elle ne s'associe pas à ses partenaires, contrairement à ce qui était attendu d'après le schéma d'interaction issu de la littérature. Il a donc été décidé de se concentrer sur la forme eIF3bC1, stable lorsqu'elle est produite seule, pour la formation de complexes.

2 Les complexes à deux partenaires

Le bilan de ces purifications est réalisé dans le **tableau 3.3**. Le bénéfice apporté par la coexpression est très net. Alors que les essais de purification des complexes par co-cassage de bactérie n'ont permis d'obtenir que 0,5mg de complexe purifié, la coexpression améliore considérablement le rendement avec en moyenne 5mg de complexe pur, facilitant dès lors les essais de cristallisation. De plus, l'effet synergique lié à l'utilisation d'un tel système est bien représenté par les purifications d'eIF3a-3cC1 qui permet de détecter eIF3a sous forme soluble. Néanmoins, là encore, la protéolyse de la sous-unité eIF3a limite fortement l'isolement d'un complexe homogène.

Ces purifications nous ont permis de confirmer l'existence des complexes dimériques envisageables d'après le schéma d'interaction issu de la littérature. Notamment, l'absence du domaine d'interaction d'eIF3b avec eIF3a dans la forme tronquée eIF3bC1 ne permet pas d'obtenir le dimère eIF3a-bC1.

3 Les complexes trimériques et tétramériques

Les purifications des sous-complexes à trois et quatre partenaires sont résumées dans le **tableau 3.4**. Un sous-complexe formé des sous-unités eIF3a, eIF3b et eIF3c représenterait le cœur fonctionnel d'eIF3, et suffit pour assurer les rôles stimulateurs de la liaison du complexe ternaire et de l'ARNm au ribosome ([Phan et al., 2001](#)). Dans ce contexte, le sous-complexe eIF3b-3i-3g, inactif dans la traduction *in vitro*, pourrait avoir un autre rôle. De façon satisfaisante, le complexe eIF3bC1-i-gC1 a pu être purifié jusqu'à homogénéité avec un bon rendement. Par contre, la forte protéolyse de la protéine eIF3a ne nous a pas permis d'isoler un complexe eIF3a-bC1-cC1 suffisamment homogène. En outre, au cours d'une purification

comportant les cinq étapes de purification le pic issu de l'étape d'interaction hydrophobe a précipité sous forme filamentaire. Ces filaments ne se dissolvent pas, même à haute force ionique et contiennent bien la protéine comme le révèle l'analyse SDS-PAGE du dépôt d'un fragment de filament.

La mise en évidence du tétramère eIF3a-cC1-i-gC1 apparaît également intéressante. En effet, d'après la littérature, il n'a pu être identifié aucune interaction entre les sous-unités eIF3i et eIF3g, d'une part, et les sous-unités eIF3a et eIF3c, d'autre part. Il apparaît que la sous-unité eIF3b est nécessaire pour ponter l'interaction entre ces deux groupes de partenaires. Nous avons cependant été en mesure d'isoler un tétramère sans la sous-unité eIF3b. La tentative de purification d'un dimère eIF3cC1-3i (**Tableau 3.3**) a finalement révélé l'absence d'interaction entre ces deux sous-unités. En revanche, les tentatives de purification du trimère eIF3cC1-3i-3gC1 ont montré qu'une faible quantité d'eIF3cC1 co-élue avec les sous-unités eIF3i et eIF3gC1. Ainsi, l'existence du complexe eIF3a-cC1-i-gC1 pourrait indiquer un effet additif lié à la protéine eIF3a. Des tentatives de purifications de complexes contenant eIF3a, eIF3g et eIF3i permettraient de préciser un nouveau réseau potentiel d'interaction, notamment en utilisant une seule protéine fusionnée à une étiquette.


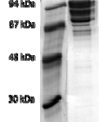

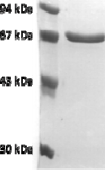

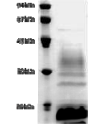

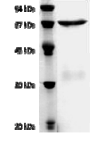

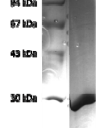

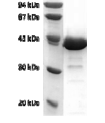
Sous-unités isolée	Purification					Quantité (mg)	Analyse SDS-PAGE	Tentatives de cristallisation	Remarques
	IMAC	DET	AEC	TM	HIC				
eIF3a	X								Aucune protéine accrochée sur résine TALON
eIF3aN1, eIF3aN2 	X		X	X		5		X	Forte protéolyse dans la zone 67-110kDa Séparation des bandes sur AEC et TM mais dégradation rapide
eIF3bC1 	X	(X)	X	X		5		X	Deux pics sur TM dont les migrations sur gel SDS-PAGE sont identiques Ils correspondent à un équilibre de dimérisation, réversible par ajout de Triton X-100 Le retrait de l'étiquette ne modifie pas l'équilibre Cristallisation essayée en présence et absence de tensioactif
eIF3bC2 	X		X	X		1,5		X	Protéine provenant d'une tentative de purification d'un dimère 3bC2-3cC1. Les protéines sont séparées par l'étape de TM Aucune association avec eIF3i ou eIF3g
eIF3cC1 	X		X	X		0,3		X	Protéine provenant d'une tentative de purification d'un dimère 3bC2-3cC1. Les protéines sont séparées par l'étape de TM
eIF3gC1 	X		X	X		6		X	L'étape AEC est réalisée à 500mM NaCl pour limiter la présence d'acides nucléiques certainement liés au domaine RRM
eIF3i 	X	(X)	X	X	(X)	5		X	Voir Partie D

Tableau 3.2 : Bilan des purifications des sous-unités isolées. Les étapes de purification correspondent aux purifications par résine TALON (IMAC), par suppression de l'étiquette His₆ (DET), par résine Q-sépharose (AEC), par tamisage moléculaire (TM) et par résine d'interaction hydrophobe de type C₃ (HIC) ([voir Matériels et Méthodes](#)).


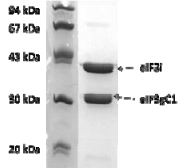

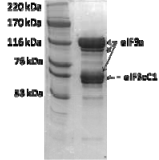

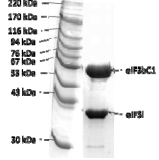

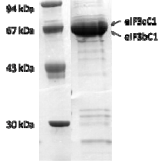

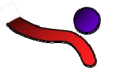
Dimère	Purification					Quantité (mg)	Analyse SDS-PAGE	Tentatives de cristallisation	Remarques
	IMAC	DET	AEC	TM	HIC				
eIF3i-gC1 	X		X	X		5		X	Deux pics sur AEC dont la migration sur gel SDS-PAGE est identique Caractérisation en masse totale (SM) n'indique pas de différences ; éventuellement présence d'un petit pic non attribuable...
eIF3a-cC1 	X		X	X					Bonne surproduction d'eIF3a mais la sous-unité subit toujours une forte protéolyse
eIF3bC1-i 	X		X	X		10		X	Deux pics sur TM dont les migrations sur gel SDS-PAGE sont identiques
eIF3bC1-cC1 	X		X	X		7		X	Protéine provenant d'une tentative de purification d'un trimère eIF3a-3bC1-3cC1.
eIF3bC1-a 	X								eIF3a n'est pas détectée
eIF3cC1-i 	X								Pas d'interaction détectable, en accord avec schéma d'interaction

Tableau 3.3 : Bilan des purifications des dimères. Les étapes de purification correspondent aux purifications par résine TALON (IMAC), par suppression de l'étiquette His₆ (DET), par résine Q-sépharose (AEC), par tamisage moléculaire (TM) et par résine d'interaction hydrophobe de type C₃ (HIC) ([voir Matériels et Méthodes](#)).

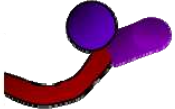
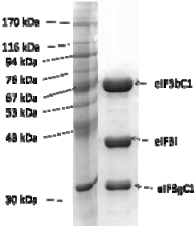

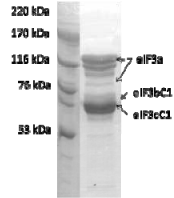
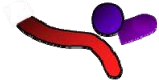

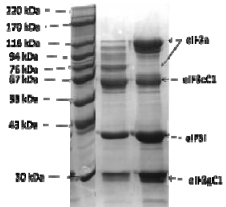

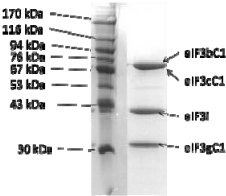
Complexe	Purification					Quantité (mg)	Analyse SDS-PAGE	Tentatives de cristallisation	Remarques
	IMAC	DET	AEC	TM	HIC				
eIF3bC1-i-gC1 	X		X	X		5		X	Deux conformations du trimère observées sur TM (voir Partie C)
eIF3a/aN1/aN2-bC1-cC1 	X	(X)	X	X	(X)				Bonne surproduction d'eIF3a mais la sous-unité subit toujours une forte protéolyse Pas de meilleurs résultats avec les tronçatures eIF3aN1 ou eIF3aN2 La protéolyse ménagée n'indique pas de fragments stables pour eIF3a
eIF3cC1-i-gC1 	X								Une faible quantité d'eIF3cC1 coélue sur résine IMAC avec eIF3i-gC1 Pas de meilleure stœchiométrie en produisant eIF3cC1 à partir d'un vecteur de type pET
eIF3a-cC1-i-gC1 	X		X	X				X	Problèmes de protéolyse d'eIF3a Stœchiométrie du complexe non satisfaisante.
eIF3bC1-cC1-i-gC1 	X	(X)	X	X	(X)	10		X	Voir Partie C

Tableau 3.4 : Bilan des purifications des trimères et tétramères. Les étapes de purification correspondent aux purifications par résine TALON (IMAC), par suppression de l'étiquette His₆ (DET), par résine Q-sépharose (AEC), par tamisage moléculaire (TM) et par résine d'interaction hydrophobe de type C₃ (HIC) ([voir Matériels et Méthodes](#)).

III Conclusion

Le développement d'une stratégie de purification basée sur un système de coexpression nous a permis de considérablement améliorer les rendements des purifications tout en limitant les opérations menant à la reconstitution de sous-complexes. Les protéines eIF3bC1, gC1, cC1, i ont pu être purifiées à homogénéité. Les hétérodimères eIF3i-gC1, eIF3bC1-i et eIF3bC1-cC1 ont également pu être purifiés à homogénéité. Toutes ces protéines et tous ces complexes ont été utilisés pour des tentatives de cristallisation. Seule la sous-unité eIF3i a cristallisé ([voir Partie D](#)). Finalement, la purification du complexe eIF3bC1-cC1-igC1 ouvre maintenant la voie à la résolution de la structure du facteur eIF3 par microscopie électronique.

PARTIE C LE COMPLEXE EIF3BC1-CC1-I-GC1, PREMIER PAS VERS LA RESOLUTION DE LA STRUCTURE DU FACTEUR EIF3

I Purification du complexe eIF3bC1-cC1-i-gC1

L'obtention de ce complexe s'est rapidement montrée intéressante puisqu'elle permet de contourner le problème de dégradation constaté pour la sous-unité eIF3a. De plus, ce complexe s'avère être bien produit et a pu être obtenu avec un bon degré de pureté (**Figure 3.12**). La purification comporte quatre étapes. Elle comprend un accrochage sur résine TALON, une étape d'échange d'ions sur résine Q-sépharose, suivis de deux étapes de polissage, tamisage puis chromatographie d'interaction hydrophobe. Il est observé lors de la purification sur tamis moléculaire l'existence d'une dimérisation partielle stable puisque l'application d'un des pics sur un nouveau tamis moléculaire ne génère qu'un pic unique. A la suite des premières tentatives de cristallisation étant restées infructueuses, l'étape d'interaction hydrophobe a été ajoutée afin de parfaire la purification. La comigration des bandes eIF3bC1 et eIF3cC1 a été confirmée par spectrométrie de masse et par des purifications de complexes ne comportant qu'une seule de ces sous-unités (**Tableau 3.3, dimères eIF3a-cC1 et eIF3bC1-3i**).

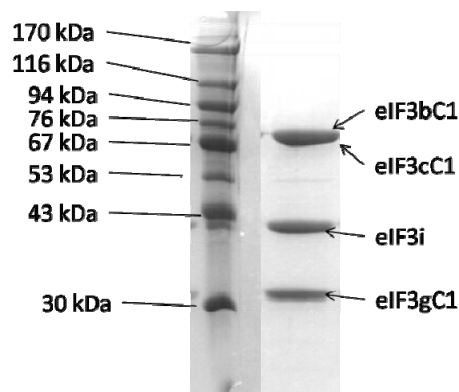


Figure 3.12 : Analyse par électrophorèse (SDS-PAGE 12%) du complexe eIF3bC1-cC1-i-gC1 purifié (piste de droite). La piste de gauche correspond à un marqueur de poids moléculaires.

Ce complexe hétérotétramérique a été caractérisé par la technique de la diffusion statique de la lumière (DLS : Diffusion Light Scattering, DynaPro Titan, Wyatt Technology). Il faut noter que tous les autres complexes obtenus ont également été caractérisés par DLS.

La technique consiste à faire traverser l'échantillon par un rayonnement laser. L'interaction onde-matière induit qu'une partie du rayonnement incident est diffusée. La diffusion dynamique de la lumière consiste à enregistrer cette partie diffusée à un angle donné (ici 90°). Le mouvement brownien des molécules de l'échantillon induit des interférences constructives ou destructives des rayonnements diffusés, créant des fluctuations de la lumière diffusée. La fonction alors analysée n'est pas la fluctuation de l'intensité moyenne au cours du temps mais la fonction d'autocorrélation. Cette fonction vise à corréliser l'état du système à un temps t donné avec ce même système à un autre temps $t+\Delta t$. Si l'espace Δt est suffisamment court, il apparaît assez intuitif que les systèmes sont fortement corrélés et plus cet espace de temps augmente et moins les systèmes seront corrélés. En fait, pour une solution monodisperse (une seule entité)

cette fonction d'autocorrélation prend la forme d'une exponentielle décroissante dont la constante de temps est liée à la taille de la particule (**Figure 3.13**). En effet, la vitesse de diffusion des molécules liée au mouvement brownien dépend de la taille des particules, tous les autres paramètres étant constants. Plus les molécules sont grandes, plus le mouvement brownien est lent et plus le signal change lentement, et donc plus la corrélation persiste longtemps. En revanche, pour des petites particules qui bougent plus rapidement, la corrélation est réduite plus rapidement. Ainsi, cette technique permet d'évaluer le rayon hydrodynamique des particules, c'est-à-dire le rayon de la sphère qui diffuserait la lumière de la même façon que la molécule. En revanche, cette technique ne fournit qu'une estimation de la masse molaire de la molécule à la différence de la diffusion statique de la lumière. En effet, la détermination de la masse molaire requiert de mesurer la dépendance angulaire de l'intensité moyenne diffusée.

Cette mesure permet en plus d'estimer la composition de l'échantillon par l'évaluation du paramètre de polydispersité. En effet, à partir de la fonction d'autocorrélation sont déduits les rayons hydrodynamiques des espèces qui composent l'échantillon. Il va en fait correspondre à chaque espèce une répartition gaussienne du rayon hydrodynamique qui représente l'homogénéité de l'espèce considérée. Notamment, si un dimère et un monomère sont présents, les pics correspondant ne pourront être résolus mais le pic résultant voit sa largeur à mi-hauteur augmentée, comparé à celui d'une espèce unique. Il a été estimé qu'un indice de polydispersité en-dessous de 15-20% correspond à une solution monodisperse. Ce critère est notamment intéressant pour la cristallogenèse, qui est facilitée pour des solutions monodisperses.

Dans notre cas, la DLS révèle bien une différence de rayons hydrodynamiques entre les espèces monomériques et dimériques identifiées par tamisage moléculaire. De plus, la masse molaire estimée du monomère est proche de la masse calculée du complexe (210 kDa) et l'indice de polydispersité est satisfaisant. La distribution plus large du dimère pourrait indiquer une contamination par le monomère (**Figure 3.13**).

De façon remarquable, une dimérisation ayant des caractéristiques similaires a été observée lors de la purification de l'hétérotrimère eIF3bC1-i-gC1. Or, une dimérisation dynamique à l'échelle de temps de la chromatographie est observée lors de la purification de la sous-unité eIF3bC1 seule. Ces observations nous amènent à proposer que la sous-unité eIF3bC1 ne fasse que nucléer la dimérisation de l'hétérotrimère et de l'hétérotétramère. La dimérisation à l'échelle de la chromatographie observée pour les complexes eIF3bC1-i-gC1 et eIF3bC1-cC1-i-gC1 serait alors induite par la sous-unité eIF3i. En effet, le repliement des β -propeller produit au centre de la structure une poche de liaison pour des peptides. Ainsi, le rapprochement des sous-unités eIF3i au sein de ces formes pourrait induire une interaction stabilisant ces assemblages. Dans cette hypothèse, l'ordre d'assemblage des sous-unités ou encore une microhétérogénéité (microprotéolyse) pourrait induire deux formes des complexes eIF3bC1-i-gC1 et eIF3bC1-cC1-i-gC1. Les complexes se trouveraient ainsi piégés irréversiblement sous l'une ou l'autre forme.

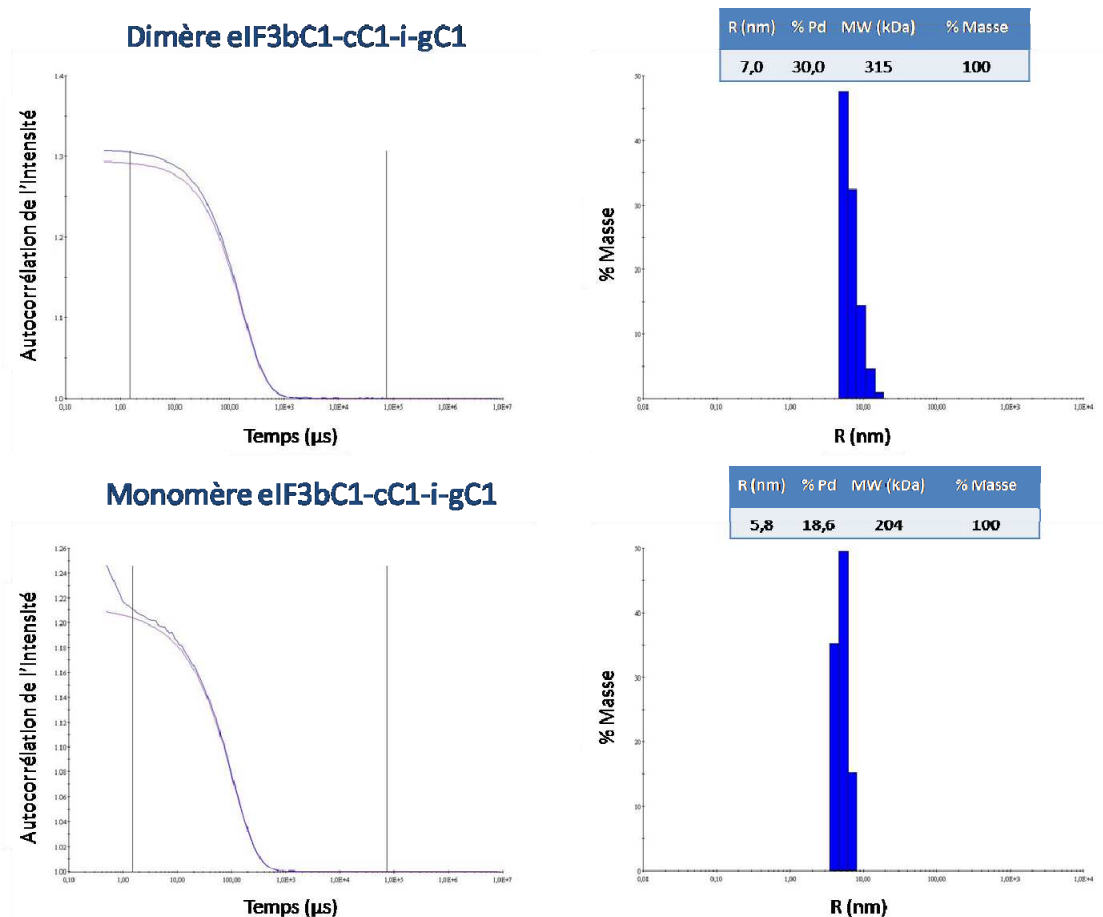


Figure 3.13 : Analyse des deux pics de l'hétérotétramère eIF3bC1-cC1-i-gC1. Les courbes d'autocorrélation correspondent aux graphiques de gauche. Courbe bleue : mesure expérimentale. Courbe violette : fonction régularisée. Les analyses des fonctions d'autocorrélation sont reportées dans les graphiques de droite et sous forme chiffrée dans les tableaux (Pd : polydispersité ; R : rayon hydrodynamique, MW : masse molaire approximative, déduite de R).

II Structure d'eIF3bC1-cC1-i-gC1 par microscopie électronique

Les difficultés à obtenir des cristaux des différents complexes nous ont incités à nous diriger vers la technique de microscopie électronique. La technique permet en effet de déterminer la structure de protéines en solution à partir du moment où leur taille est suffisamment importante (supérieure à 200-250kDa). Le complexe eIF3bC1-cC1-i-gC1 a une masse molaire de 210 kDa, acceptable pour tenter de résoudre sa structure par cette technique. La limitation principale de cette technique reste la résolution obtenue qui est conditionnée par la mobilité des espèces observées et les imperfections du système optique. La technique repose sur l'observation dans l'espace direct des objets, à l'instar de la microscopie optique. La microscopie électronique utilise non pas des photons mais des électrons accélérés qui vont interagir avec l'échantillon. Le faisceau transmis est enregistré, l'échantillon doit donc être suffisamment mince pour permettre aux électrons de le traverser. Un système de lentilles magnétiques permet de projeter l'image de l'échantillon sur un écran fluorescent qui transforme l'image électronique en image optique.

L'échantillon étudié est déposé sur une grille de carbone préalablement nettoyée dans un plasma. Le surplus est absorbé puis un agent contrastant est ajouté. Ceci est nécessaire dans notre cas pour améliorer le contraste très limité sur des objets dont la masse est proche de la masse limite. De plus, le vide poussé imposé dans la colonne requiert d'insérer un échantillon solide. Communément, l'agent contrastant est l'acétate d'uranyle. Une goutte est apposée sur la grille sur laquelle l'échantillon est adsorbé. Le surplus est enlevé puis la grille est mise à sécher. En séchant, l'agent contrastant prend la forme des objets qu'il recouvre et c'est en fait cette empreinte qui est analysée. Pour l'étude d'objets moléculaires plus gros, la technique de cryomicroscopie électronique est utilisée. Dans ce cas, l'échantillon est de la même façon adsorbé sur une grille de carbone qui est plongée rapidement dans l'éthane liquide afin de piéger les molécules. La grille est alors insérée dans un porte échantillon spécifique avant d'être analysée. La résolution d'une structure par microscopie électronique nécessite l'acquisition d'un grand nombre d'images. La structure obtenue résulte alors d'une moyenne sur un grand nombre de particules identifiées. Des logiciels informatiques organisent ensuite ces images par classes, qui correspondent à différentes orientations des molécules étudiées. Le recoupement de ces différentes vues permet d'obtenir l'image tridimensionnelle de l'objet. Il est ainsi nécessaire que l'objet ait une forme suffisamment dissymétrique afin de pouvoir différencier ses différentes orientations.

Des images du complexe eIF3bC1-cC1-i-gC1 par microscopie électronique ont été réalisées par Pierre Damien Coureux, enseignant chercheur au laboratoire. Ces enregistrements ont été réalisés à l'Institut de Minéralogie et physique des Milieux Condensés (Paris, Campus Boucicaut) sur un microscope Jeol 2100. La structure est actuellement en cours de résolution, les résultats préliminaires sont présentés dans la **Figure 3.14**.

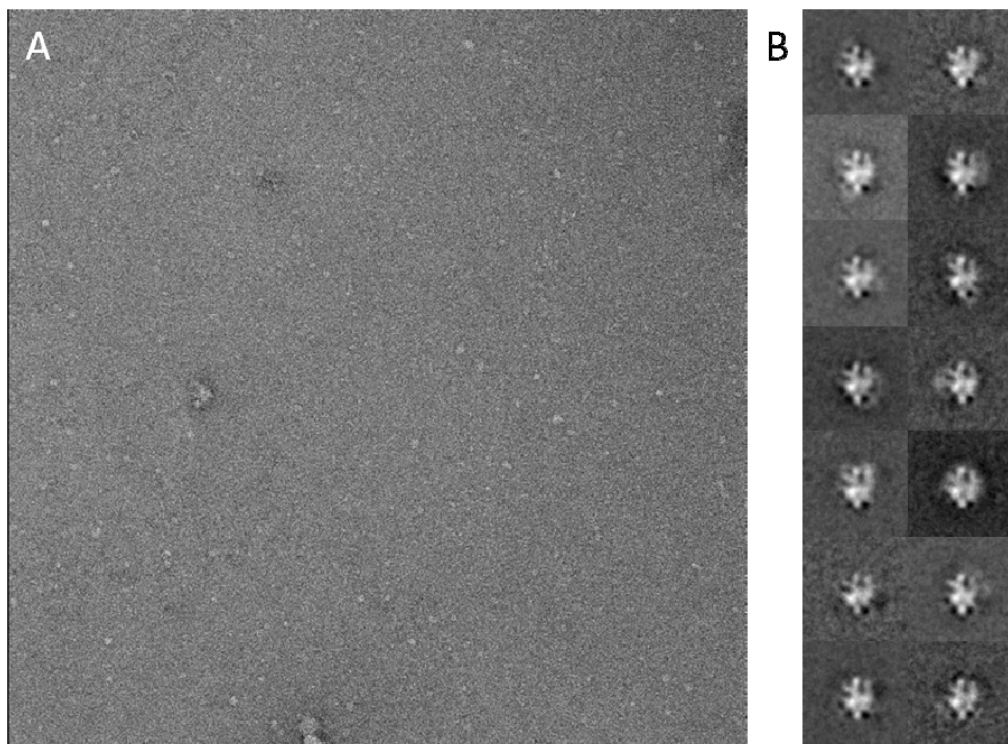


Figure 3.14 : Images de microscopie électronique de l'hétérotétramère eIF3bC1-cC1-i-gC1. (A) Image réalisée en coloration négative (acétate d'uranyle 2%) à un grossissement de 65000 sur un microscope Jeol2100 (FEG 200kV). L'image est enregistrée à l'aide d'une caméra 4K. (B) Moyennes de classes obtenues à partir de 3000 particules sélectionnées.

III Vers une structure du facteur eIF3 par microscopie électronique

La continuité de la résolution de cette structure réside dans l'attribution de la densité électronique aux différentes sous-unités, ce qui ne peut être envisagé que dans un système où l'on peut aisément produire des sous-complexes. En effet, si l'on dispose des structures d'un complexe et d'un sous-complexe où il manque une sous-unité, la différence entre ces deux cartes fournit la densité et la localisation de la sous-unité manquante. Les résultats sur les purifications indiquent que les complexes eIF3bC1-cC1-i et eIF3bC1-cC1-gC1 devrait être facilement produits et purifiés. Leurs masses, 182kDa et 173kDa respectivement, restent acceptables pour espérer résoudre leurs structures. Nous pourrions ainsi localiser les sous-unités eIF3i et eIF3gC1 au sein du complexe.

L'obtention du complexe pentamérique constitue le point majeur du projet. En outre, sa taille (323 kDa) est adaptée aux études par microscopie électronique. Dans les essais réalisés jusqu'à présent, en particulier en utilisant la stratégie de coexpression, la stœchiométrie du complexe apparaît un peu décevante en ce qui concerne eIF3a. Cela est dû notamment à la présence de nombreuses bandes de protéolyse sous celle d'eIF3a. Il est à noter que l'absence de ces bandes lors des premiers essais doit être liée à une purification plus rapide.

Une autre difficulté provient de la comigration des bandes d'eIF3bC1 et d'eIF3cC1, mise en évidence par l'analyse en SM, qui limite l'évaluation de la stœchiométrie. La diminution du pourcentage d'acrylamide des gels de migration ne permet pas de résoudre les deux bandes. Consécutivement, le produit de la purification de la co-transformation de bactéries par les plasmides pET32b-3bC1, pCDF-3a-3cC1 et pCOLA-3i-3gC1 a été analysé par un gel bidimensionnel. Les protéines sont d'abord séparées en fonction de leurs points isoélectriques (gamme théorique des différentes sous-unités : 5-7 ; gamme de la bande tampon : 4-7) puis en fonction de leurs masses. Des spots correspondants aux protéines 3i, 3gC1 sont clairement distinguables. Dans la zone des poids moléculaires correspondant à eIF3bC1 et eIF3cC1, plusieurs spots sont repérés ; en augmentant la quantité de l'échantillon appliqué, deux groupes apparaissent. Les points isoélectriques d'eIF3bC1 et d'eIF3cC1 étant respectivement de 6,5 et 5,9, l'attribution des spots peut être prédite. De plus, il est attendu une différence de masse (masses molaires attendues : eIF3bC1 67kDa et eIF3cC1 77kDa), qui est effectivement détectable malgré des migrations très proches, confirmant ainsi la présence des deux protéines. La sous-unité eIF3a n'est pas identifiable sur ces gels. En revanche, des protéines semblent présentes sur la bande tampon, lorsqu'elle est colorée au bleu de coomassie, indiquant que la protéine a certainement dû précipiter au cours de la focalisation. L'évaluation de la stœchiométrie par cette analyse paraît difficile à cause des dédoublements de spots. Notamment, eIF3i semble être résolue sous la forme de trois spots. La présence des différents spots doit être confirmée en réalisant de nouvelles migrations bidimensionnelles pour s'affranchir d'artefacts pouvant survenir au moment de la focalisation. Si ceux-ci sont à nouveau identifiés, nous prévoyons de caractériser ces isoformes potentielles.

Le bilan de ces purifications semble indiquer un défaut d'expression des protéines eIF3aN1 et eIF3aN2 lorsqu'elles sont coexprimées avec plus de trois partenaires. En outre, les deux systèmes de coexpression envisagés ne permettent pas d'obtenir un complexe stœchiométrique, que ce soit avec la forme native d'eIF3a, eIF3aN1 ou encore eIF3aN2. L'utilisation d'eIF3aN1 ou eIF3aN2 en tant que seul partenaire du complexe fusionné à une étiquette ne permet d'obtenir en sortie de résine IMAC qu'une quantité de complexe très faible pour procéder à d'autres étapes chromatographiques

alors que toute la sous-unité ne semble pas accrochée à la résine. Cela indique un masquage possible de l'étiquette au sein du complexe.

Une observation confirme cette hypothèse. Dans le complexe eIF3a-3bC1-3cC1, eIF3a et eIF3bC1 portent une étiquette polyhistidine. L'étiquette d'eIF3bC1 peut être enlevée par la thrombine. Par contre, la sous-unité eIF3a produite à partir de la construction utilisée (pCDF-3a-3cC1) ne comporte pas de site de coupure. Le complexe eIF3a-bC1-cC1 hydrolysé par la thrombine comporte donc toujours une étiquette fusionnée à eIF3a. Or, celui-ci n'est plus retenu sur la résine TALON. Ceci indique un enfouissement de l'étiquette portée par cette sous-unité. De plus, la présence de l'étiquette a pu être confirmée par l'analyse de séquençage N-terminal, réalisée sur l'hétérotrimère hydrolysé, qui indique que la protéine eIF3a possède effectivement toujours la partie polyhistidine. Finalement, la purification d'un complexe eIF3a-3cC1 par accrochage sur résine IMAC indique par comparaison avec la purification de l'hétérotrimère eIF3a-bC1-cC1 que l'enfouissement est probablement induit par l'ajout de la sous-unité eIF3bC1 au complexe.

Dans un autre cas, il a été essayé d'ajouter une étiquette à eIF3cC1 mais très peu de complexe est retenu, indiquant une probable excision de l'étiquette fusionnée à l'extrémité N-terminale d'eIF3cC1 (région de protéolyse identifiée par SM) ou alors un masquage de l'étiquette au sein du pentamère.

Dans ce contexte, nous avons envisagé une reconstitution du complexe à partir de la sous-unité eIF3aN1 ou eIF3aN2, d'une part, et du tétramère eIF3bC1-3cC1-3i-3gC1, d'autre part, facilement purifié et stable. Un essai préliminaire a été réalisé en mélangeant des culots de bactéries surexprimant le tétramère au sein duquel aucun partenaire n'est fusionné à une étiquette d'une part et une des formes d'eIF3a depuis un vecteur de type pET15b. La forte surproduction d'eIF3a a permis de fixer l'ensemble du tétramère. En revanche, le large excès d'eIF3a est difficilement éliminable. D'une part, la forte affinité d'eIF3a sur la résine échangeuse d'ions impose une élution en réalisant un saut brutal de concentration en sels. D'autre part, son poids moléculaire élevé impose plusieurs tamisages moléculaires pour sélectionner le complexe, mais cela favorise sa protéolyse.

La stabilité du tétramère au regard de celle d'eIF3a nous indique une autre piste consistant à le purifier extemporanément, l'accrocher sur la résine IMAC puis appliquer un lysat de cellules surproduisant une forme de la sous-unité eIF3a sans étiquette. Enfin, nous avons prévu de produire la sous-unité eIF3a (et ses différentes formes) fusionnée à une étiquette à son extrémité C-terminale. La protéolyse survenant par cette partie, l'étiquetage C-terminal devrait nous permettre de sélectionner la protéine intacte par accrochage sur résine IMAC. Cette construction nous permettra également peut-être d'obtenir les sous-complexes comportant la sous-unité eIF3a intacte, facilitant dès lors la résolution de la structure par microscopie électronique.

PARTIE D LES CRISTAUX D'EIF3I DE *SACCHAROMYCES CERVISIAE* :

UN PROBLEME CRISTALLOGRAPHIQUE

I Purification de la protéine

Cette sous-unité s'est révélée dès les premiers tests d'expression être fortement surexprimée (50-100mg de protéine éluée de la résine TALON à partir d'un litre de culture) et très stable. Le protocole de purification a du être optimisé afin d'obtenir la cristallisation de la protéine. Brièvement, un accrochage sur résine TALON est suivi d'une étape de suppression de l'étiquette polyhistidine et séparation des différents produits de la coupure. La protéine étant instable à basse force ionique, le pic collecté correspondant à la protéine sans étiquette est dilué dans un tampon sans sel puis rapidement appliqué sur une résine échangeuse d'anions. Enfin, un polissage est réalisé sur tamis moléculaire. L'étape d'échange d'ions génère deux pics dont la hauteur varie en fonction des purifications. L'analyse sur SDS-PAGE ne fait pas apparaître de différences. Les analyses en SM n'ont pu amener d'argument convaincant. La cristallisation n'est obtenue que lorsque la protéine est dénuée d'étiquette et provient du premier pic d'élution de la résine échangeuse d'ions.

II Principe de la cristallisation

1 Principe

Comme tout processus de cristallisation, la « biocristallisation » est générée par les trois étapes de nucléation, croissance du cristal et arrêt de la croissance, et influencée par de multiples paramètres. La difficulté dans la cristallisation de macromolécules d'origine biologique repose essentiellement sur le nombre plus important de paramètres limitant le processus (abondance, pureté...) mais surtout sur les caractéristiques physico-chimiques particulières des protéines. Notamment, la stabilité des protéines est bien souvent restreinte à une gamme relativement étroite de température et de pH. La complexité du système est renforcée par la flexibilité dans la conformation des protéines et leurs éventuelles modifications chimiques, induisant une hétérogénéité dans la solution et gênant alors la cristallisation. La pureté de l'échantillon est en effet un facteur important afin d'assurer une croissance cristalline optimale. La mise au point d'une stratégie de purification adaptée à la cristallisation apparaît dès lors comme la première étape dans cette démarche.

L'obtention de cristaux de toute molécule requiert de placer le composé dans un état thermodynamiquement instable dit sursaturé. Le retour à une phase d'équilibre peut alors évoluer vers la formation d'une phase amorphe (identifiée comme une précipitation) ou bien conduire à la croissance d'une phase cristalline. L'obtention de la sursaturation peut être obtenue par simple évaporation du solvant mais cette technique n'est pas adaptée aux objets macromoléculaires sensibles que sont les protéines. La démarche consiste alors à une modification lente du solvant dans lequel la protéine est diluée, notamment par l'ajout lent d'un agent dit précipitant ou bien en modifiant les conditions physiques comme le pH. Les agents précipitants généralement rencontrés peuvent être des sels (NaCl, sulfate d'ammonium...), des polymères (éthylène glycol, PEG...), certains alcools (hexanediol ou méthylpentanediol) ou encore des composés organiques non volatils. Dans tous les cas, le but est d'atteindre des concentrations en protéine et en agent précipitant

telles que les deux composants se retrouvent en compétition pour la fixation de molécules d'eau assurant leurs solvatations. La combinaison des effets du pH et de l'agent précipitant est souvent mise à profit pour obtenir la cristallisation.

La mise en œuvre expérimentale généralement utilisée repose sur la technique de diffusion de vapeur. Il peut-être également procédé à la cristallisation par diffusion de liquide, en gel, par changement de température ou encore par ensemencement. La technique de diffusion de vapeur connaît deux variantes assez proches. La première consiste à déposer une goutte sur un support dans une enceinte close contenant la condition de cristallisation, plus communément appelée goutte assise. Alternativement, la goutte peut-être déposée sur une lamelle en verre retournée sur un réservoir scellé contenant la condition de cristallisation, la goutte est dite suspendue. La goutte déposée (1 à 20 μ L) est un mélange de la liqueur mère, à savoir la protéine pure et concentrée (en général de 2 à 15mg/mL), et de la solution de cristallisation dans des proportions pouvant varier, communément 1:1. Le réservoir contient un volume de solution de cristallisation grand devant celui de la goutte (100 à 500 μ L). Le système totalement étanche permet d'éviter le phénomène d'évaporation du solvant vers l'environnement. Ainsi, les seuls transferts de matières permis vont concerner l'eau ou les composés volatils présents. Dans des conditions usuelles (ratio solution de protéine/condition de cristallisation 1:1), la concentration de la goutte en agent précipitant est la moitié de celle du réservoir. L'eau, dans le cas des sels et des polymères, va alors s'échanger depuis la goutte vers le réservoir afin d'assurer un équilibre des concentrations en agent précipitant dans la goutte et dans le réservoir. Consécutivement, le volume de la goutte diminue. Dans le cas où la condition de cristallisation contient également (ou exclusivement) un composé volatile, le transfert de ces composants se fera depuis le réservoir vers la goutte afin de permettre d'atteindre l'équilibre. Cette technique étant basée sur une diffusion, le temps pris pour atteindre l'équilibre permet l'organisation des macromolécules aboutissant à la formation d'un réseau cristallin.

2 Cristallisation d'eIF3i

Nous disposons au laboratoire d'un robot Genesis RSP100 (TECAN) adapté à la réalisation de tests de cristallisation. A partir de matrices de conditions de cristallisation de 96 puits commerciales ou préparées au laboratoire, le robot distribue les solutions de cristallisation dans les puits de boîtes spécialement destinées à la cristallisation par diffusion de vapeur en gouttes assises. Les gouttes sont ensuite réalisées en mélangeant 1 μ L de liqueur mère et 1 μ L de condition de cristallisation. Dans le cas d'eIF3i, la cristallisation a été obtenue avec la matrice SaltRX (Hampton Research) dans la condition 58 [1,8M Na/KPO₄ pH8,2] après 3 semaines à 4°C. Les cristaux se présentaient sous la forme de très fines aiguilles, souvent organisées tels des oursins (**Figure 3.15(A)**).

L'étape suivant cette première cristallisation consiste à les reproduire et surtout les optimiser afin d'obtenir des cristaux de tailles suffisantes pour être soumis à un rayonnement X. La répétition de la cristallisation a été effectuée d'une part dans des boîtes à 24 puits, permettant de réaliser des gouttes à la main avec de plus grands volumes, en réalisant une gamme en agent précipitant et en pH (gouttes de 2+2 μ L/volume de condition de cristallisation du réservoir : 500 μ L). D'autre part, nous avons utilisé des matrices d'additifs comportant des composés pouvant permettre d'améliorer la taille et/ou le pouvoir diffractant des cristaux. Les gouttes ont alors été réalisées à la main dans une boîte à 96 puits du même type que celles utilisées par le robot. De façon surprenante, les cristaux n'ont pu être reproduits en confectionnant les gouttes manuellement en boîte 24 puits. Au contraire, après

3 semaines à 4°C, la boîte de tests d'additifs comportait plusieurs cristaux ayant la forme de baguettes de tailles bien plus importantes (200x50x20µm) (**Figure 3.15(B)**). De plus, il apparaît qu'une répétition de la condition de cristallisation sans additif dans cette même boîte a permis d'obtenir des cristaux de tailles similaires à ceux obtenus avec les meilleurs additifs. Cela a indiqué une forte dépendance de la taille et la forme des cristaux en fonction de la manière de réaliser les gouttes plutôt qu'une influence des additifs.

Les cristallisations suivantes ont confirmé que des cristaux de taille acceptable peuvent être obtenus dans les boîtes de 96 puits par diffusion de vapeur en gouttes assises si la goutte est faite selon la technique suivante. 1µL de liqueur mère est déposé dans le fond du puits. 1µL de la condition de cristallisation est ensuite fait perler au bout d'un cône puis apposé sur la protéine. Il semble en effet que lorsque la protéine et la solution de cristallisation sont mélangées avant de confectionner la goutte, seuls des cristaux sous forme d'oursins sont obtenus. En outre, les réalisations de gammes de concentrations en agent précipitant et en pH dans ce système n'ont pas révélé de meilleures conditions. La cristallisation est de plus peu reproductible et de nombreuses répétitions sont nécessaires afin d'obtenir des cristaux de tailles suffisantes.

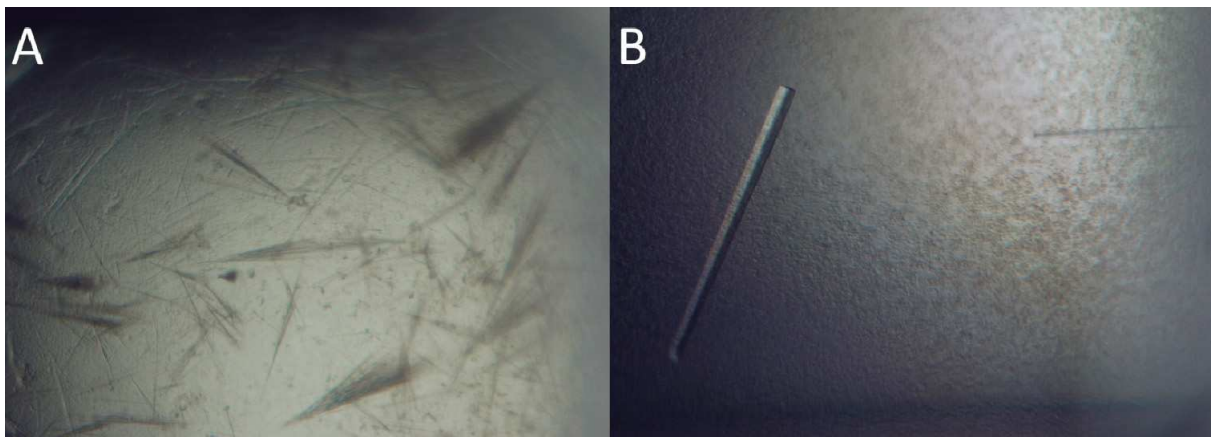


Figure 3.15 : Cristaux d'eIF3i obtenus dans la condition 1,8M Na/KPO₄ pH8,2. La photo A représente des cristaux sous la forme de fine baguettes. La photo B illustre la croissance de ces cristaux lorsqu'ils sont réalisés selon la méthode décrite dans cette partie.

III Diffraction par les rayons X et acquisition des données

1 La diffraction des rayons X par les cristaux et conditions de diffraction

Un cristal est défini par une grande organisation liée à sa périodicité. Lorsqu'un rayonnement électromagnétique tels les rayons X rencontrent un cristal, une partie du rayonnement est absorbée par l'échantillon alors qu'une autre partie, de même longueur d'onde que le rayon incident, est diffusée dans toutes les directions de l'espace. Or, les distances interatomiques au sein des macromolécules qui composent le cristal, de l'ordre de 1,5Å, sont du même ordre de grandeur que la longueur d'onde du rayonnement X. Dans ce cas, les rayonnements diffusés par les différents atomes du réseau vont produire des interférences alternativement constructives et destructives. Ainsi, selon la direction de l'espace, les flux de photons X seront plus ou moins importants, créant alors le phénomène de diffraction.

Les directions dans lesquelles les interférences sont constructives, appelées taches ou pics de diffraction, peuvent être déterminées simplement par la loi de Bragg :

$$2d \sin \theta = n \times \lambda$$

Où :

d = distance interréticulaire, c'est-à-dire distance entre deux plans cristallographiques

θ = demi-angle de déviation (moitié de l'angle entre le faisceau incident et le faisceau diffracté) ;

n = ordre de réflexion (nombre entier) ;

λ = longueur d'onde des rayons X.

Si le rayonnement incident est monochromatique de longueur d'onde λ , alors l'onde diffusée en \vec{x} , à l'instant t , par un atome j peut s'écrire sous la forme :

$$\psi_j(\vec{x}, t) = \psi_0 \cdot f_j \cdot e^{i(\omega t + \varphi(\vec{x}))}$$

ω = pulsation de l'onde ($\omega = 2\pi c/\lambda$)

f_j = le facteur de diffusion atomique, qui dépend de la densité du nuage électronique de l'atome, donc de sa nature chimique

φ = le déphasage de l'onde en \vec{x} par rapport à l'origine spatiale

Le déphasage $\varphi(\vec{x})$ est la somme du déphasage entre la source et le point \vec{r}_j où est placé l'atome j et de celui entre l'atome j et le point d'observation \vec{x} . Si l'on définit le vecteur de diffraction $\vec{K} = \vec{k}' - \vec{k}$ comme la différence entre le vecteur d'onde de l'onde diffusée et celui de l'onde incidente, alors :

$$\varphi(\vec{x}) = 2\pi(\vec{K} \cdot \vec{r}_j - \vec{k}' \cdot \vec{x})$$

D'où :

$$\psi_j = \psi_0 \cdot e^{i(\omega t - 2\pi \vec{k}' \cdot \vec{x})} \cdot f_j \cdot e^{2i\pi \vec{K} \cdot \vec{r}_j}$$

Si l'on se place maintenant à l'échelle de la maille du cristal comportant n atomes, l'onde diffusée en un point de l'espace s'écrira :

$$\psi' = \sum_{j=1}^n \psi_j = \psi_0 \cdot e^{i(\omega t - 2\pi \vec{k}' \cdot \vec{x})} \cdot \sum_{j=1}^n f_j \cdot e^{2i\pi \vec{K} \cdot \vec{r}_j}$$

On définit ainsi le facteur de structure :

$$F(\vec{K}) = \sum_{j=1}^n f_j \cdot e^{2i\pi \vec{K} \cdot \vec{r}_j}$$

Nous avons considéré l'onde diffusée par un atome du cristal. Or, en toute rigueur, l'onde est diffusée par le cortège électronique, il convient alors de substituer la somme discrète des facteurs de diffusion atomique par une intégrale sur le volume v de la maille de la densité électronique $\rho(\vec{r})$. Le facteur de structure prend alors la forme :

$$F(\vec{K}) = \iiint_v \rho(\vec{r}) \cdot e^{2i\pi\vec{K}\cdot\vec{r}_j} \cdot dV$$

L'intensité diffractée en un point \vec{x} de l'espace est proportionnelle au carré de la norme du vecteur de la fonction d'onde :

$$I(\vec{x}) \propto |F(\vec{K})|^2$$

Chaque tâche de diffraction correspond ainsi à un maximum local de la norme du facteur de structure. Celle-ci sera maximale lorsque les ondes diffusées par deux atomes sont en phase, soit pour deux atomes j et l :

$$2\pi\vec{K}\cdot\vec{r}_j = 2\pi\vec{K}\cdot\vec{r}_l [2\pi] \quad (1)$$

Si l'on définit maintenant la base du réseau direct $(\vec{a}, \vec{b}, \vec{c})$, alors :

$$\vec{r}_j = x_j \cdot \vec{a} + y_j \cdot \vec{b} + z_j \cdot \vec{c}$$

Avec x_j, y_j, z_j des nombres entiers positifs inférieurs ou égaux à 1.

Considérons maintenant l'espace de diffraction, c'est-à-dire l'espace réciproque $(\vec{a}^*, \vec{b}^*, \vec{c}^*)$ défini par :

$$\vec{a}^* = \frac{\vec{b} \wedge \vec{c}}{V} \quad ; \quad \vec{b}^* = \frac{\vec{c} \wedge \vec{a}}{V} \quad ; \quad \vec{c}^* = \frac{\vec{a} \wedge \vec{b}}{V}$$

Où V est le volume de la maille : $V = \vec{a}^* \cdot (\vec{b}^* \wedge \vec{c}^*)$

La condition générale de diffraction nécessite d'appliquer la relation (1) à tous les atomes de la maille. Cela impose aux produits scalaires $\vec{K} \cdot \vec{e}_i$ ($i=1, 2, 3$ correspondant respectivement à $\vec{a}, \vec{b}, \vec{c}$) d'être des entiers. La définition de l'espace réciproque impose de plus que $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$. Il peut donc être défini dans la base réciproque une combinaison linéaire permettant de décrire le vecteur de diffraction. Ces indices, notés h, k et l , sont alors des entiers, appelés indices de Miller (hkl). Ils permettent de décrire les plans réticulaires d'un cristal.

$$\vec{K}_{hkl} = h \cdot \vec{a}^* + k \cdot \vec{b}^* + l \cdot \vec{c}^*$$

Un cliché de diffraction se présente ainsi sous la forme d'un ensemble de tâches de diffraction auxquelles sont attribuées des indices (hkl). De la même façon, le facteur de structure peut s'écrire en fonction de ces indices :

$$F_{hkl} = F(\overrightarrow{K}_{hkl}) = \iiint_V \rho(x, y, z) \cdot e^{2i\pi(hx+ky+zl)} \cdot dv$$

La densité électronique est alors obtenue par une transformée de Fourier inverse permettant d'obtenir cette fonction de l'espace réel par une somme sur l'ensemble des nœuds du réseau :

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} \cdot e^{2i\pi(hx+ky+zl)}$$

Néanmoins, l'enregistrement d'un cliché de diffraction et donc d'intensités ne permet d'obtenir que le module de F_{hkl} et l'information sur les phases est perdue, empêchant de déterminer la densité électronique. Différentes méthodes ont été mises au point afin de les calculer (MIR, MAD, remplacement moléculaire,...). Le calcul de phases représente une étape critique limitante après l'obtention de cristaux et d'une diffraction correcte.

2 Source du rayonnement X et cryoprotection

La source de rayons X utilisée conditionne la qualité de la diffraction. De nombreux efforts sont donc dirigés vers l'obtention de sources les plus monochromatiques et intenses possibles. Les sources classiques de rayons X génèrent des photons par bombardement d'une anode en métal par des électrons accélérés. En choisissant une raie d'émission de ce métal, un rayonnement monochromatique peut être utilisé pour provoquer la diffraction d'un cristal. Néanmoins, l'intensité du rayonnement reste faible et ce type de source est intéressant, à l'heure actuelle, pour tester la diffraction de cristaux et l'optimiser directement au laboratoire.

Le développement des sources synchrotron a permis le plein essor de la cristallographie, en permettant d'utiliser des cristaux de plus en plus petits et d'en acquérir les données de diffraction en un temps de plus en plus court. Le rayonnement synchrotron est émis par des électrons (ou des positrons) accélérés à une vitesse proche de celle de la lumière dans une structure circulaire comportant des électroaimants. La lumière émise par ces particules correspond à un rayonnement blanc polychromatique. Tout un jeu de monochromateurs et de lentilles sont donc installés en amont de la salle d'enregistrement afin d'obtenir un faisceau bien focalisé et monochromatique. Un avantage considérable du rayonnement synchrotron est le choix de la longueur d'onde. Combiné aux avancées dans la conception des détecteurs, le rayonnement synchrotron permet d'obtenir des données de meilleure qualité et de meilleure résolution. L'ensemble des jeux de données des cristaux acquis ont été enregistrés aux sources synchrotron de l'ESRF de Grenoble (European Synchrotron Radiation Facilities) et plus récemment à celles du Synchrotron SOLEIL de Saint-Aubin.

Le rayonnement X produit par ces sources étant très intense, des techniques permettant de cryoprotéger les cristaux se sont développées. En effet, les rayons X provoquent la formation de radicaux qui détruisent l'ordre cristallin. Le refroidissement des cristaux à une température avoisinant les 100 K par un flux d'azote permet d'augmenter la durée de vie des cristaux. Le problème de cette technique est la formation de cristaux de glace. En effet, les cristaux sont conservés dans la liqueur mère qui contient en général beaucoup d'eau. Les cristaux de protéines eux-mêmes contiennent suffisamment de solvant pour que le refroidissement rapide de la boucle, sur laquelle est maintenu le cristal dans une goutte de solution de cristallisation, permette la formation de cristaux de glace, qui

détruiraient le cristal et perturberaient la diffraction. Il doit donc être déterminé une solution cryoprotectante suffisamment proche de la solution de cristallisation pour ne pas altérer le cristal et donc sa diffraction. Généralement, des composés tels que le glycérol, l'éthylène glycol, le PEG ou le MPD sont utilisés. Ceux-ci permettent, quand ils sont suffisamment concentrés, de favoriser la formation de glace amorphe. Le choix est orienté par les résultats d'étude systématiques ([McFerrin and Snell, 2002](#); [Rubinson et al., 2000](#)) quand des données sont disponibles. Plusieurs solutions sont testées pour la diffraction afin de déterminer la solution optimale.

Dans le cas d'eIF3i, la littérature indiquait des solutions contenant de l'éthylène glycol, du glycérol ou du PEG3350. Très clairement, les tests de diffraction ont désigné le glycérol, ajouté à 20% (v/v) à la solution de cristallisation, comme le meilleur agent cryoprotectant. Il a également pu être remarqué que le transfert du cristal dans cette solution devait être le plus bref possible avant sa congélation rapide dans l'éthane ou l'azote liquides afin de préserver son pouvoir diffractant. Enfin, le malonate de sodium est décrit comme un sel, cryoprotectant à concentration élevée, adapté à la cryoprotection de protéines qui cristallisent dans des conditions contenant une haute force ionique. Ce sel a été utilisé avec succès pour améliorer les clichés de diffraction ([Holyoak et al., 2003](#)). Toutefois, dans le cas d'eIF3i, l'utilisation de malonate a plutôt provoqué une diminution sensible de la diffraction.

3 Collecte et traitement des données

Le cliché de diffraction est une image du réseau réciproque du cristal (**Figure 3.16**). L'acquisition d'une première image, ou éventuellement de deux images dans deux directions du cristal (généralement une rotation de 90° du cristal), permet dans un premier temps de déterminer le groupe d'espace du cristal et les paramètres de maille. Le programme mosflm de la suite de programmes CCP4 permet d'évaluer les paramètres de maille, le groupe d'espace, sa mosaïcité ainsi que son orientation. Or, comme nous l'avons vu, la condition générale d'observation du faisceau diffracté impose que le vecteur de diffraction coïncide avec un nœud du réseau réciproque. Pour collecter l'espace de diffraction correspondant au réseau réciproque, il est donc nécessaire d'amener chaque nœud du réseau réciproque en position de diffraction. Ceci est accompli en faisant tourner le cristal autour d'un axe de rotation perpendiculaire au faisceau de rayons X incident. Le réseau réciproque décrit alors le même mouvement que le réseau direct. Il faudra ainsi effectuer une rotation d'un certain angle minimum pour collecter des données couvrant la plus grande partie de l'espace réciproque, de manière homogène, sans laisser de trop grandes zones aveugles. Pour que les statistiques soient significatives, il faut collecter plus d'une fois chaque réflexion ou l'une de ses réflexions équivalentes.

Plus le réseau direct est grand, comme dans le cas des macromolécules biologiques, plus le réseau réciproque sera petit et un grand nombre de taches de diffraction devra donc être mesuré. La collecte des données se fait grâce à des détecteurs bidimensionnels permettant d'enregistrer des tranches de l'espace de diffraction. La rotation du cristal se fait par incrément déterminé. Celui-ci doit être choisi afin d'éviter la superposition des taches. Des taches collectées partiellement se retrouvent sur plusieurs images successives. Un paramètre d'ordre, la mosaïcité d'un cristal, est reliée à l'angle de rotation pendant lequel une réflexion est mesurable. En fonction de la méthode d'intégration utilisée, les taches entières peuvent être importantes pour la mise à l'échelle des images entre elles. Ainsi, pour une intégration avec mosflm, il est préférable d'utiliser une stratégie de collecte

donnant un nombre significatif de taches entières. Ce point est moins critique lorsqu'on utilise des logiciels d'intégration qui utilisent les profils 3D, comme XDS ou Denzo. Bien que dans la plupart des cas les collectes se fassent sans orientation préalable du cristal, il faut que l'espace de diffraction soit enregistré avec soin pour optimiser le traitement des données.

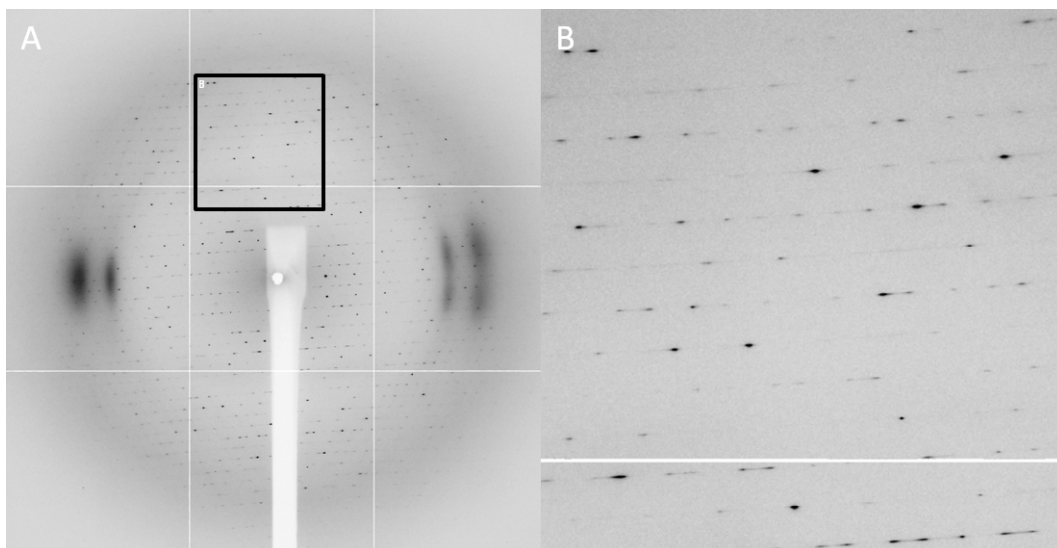


Figure 3.16 : Cliché de diffraction d'un cristal d'eIF3i natif. L'acquisition a été réalisée sur la ligne Id23eh1 de l'ESRF (Grenoble). (A) Cliché produit par une rotation de 1° du cristal exposé au rayonnement pendant 2s (transmission : 20%). (B) Vue rapprochée des taches de diffraction.

Le programme mosflm permet de prédire la zone de l'espace à balayer et le pas d'incrément de l'angle de rotation afin d'obtenir les données les plus complètes possibles. Les paramètres déterminés et notamment le groupe d'espace indiquant les zones redondantes de l'espace permettent d'évaluer la région de l'espace à enregistrer alors que les paramètres de la maille et la mosaïcité détermineront l'incrément de rotation maximal assurant une bonne résolution des tâches. Généralement, des incréments de 1° sont utilisés. Les cristaux d'eIF3i perdent néanmoins rapidement leur pouvoir diffractant au cours de l'acquisition. Il a ainsi été nécessaire de combiner des rotations (30-50°) avec des translations du cristal pour acquérir la totalité de l'espace. La translation suivant l'axe des baguettes n'altère pas la qualité des données.

Les données peuvent-être traitées alternativement par trois suites de programmes : la suite CCP4, XDS ou HKL. Les cristaux d'eIF3i sont caractérisés par une grande mosaïcité et des taches de formes allongées (**Figure 3.16(B)**). Dans ce cas, le logiciel mosflm a fourni des paramètres de mailles les plus fiables. Le traitement par XDS n'a pu être effectué qu'en injectant ces paramètres. Le programme Denzo de la suite HKL permet une détermination de la maille dans la plupart des cas. Cette détermination met en jeu un algorithme d'autoindexation ([Kabsch, 1988](#)). Des pics correspondant aux taches de diffraction sont mesurés sur une ou plusieurs images. Une liste de vecteurs entre ces pics est ensuite écrite. Les pics correspondent à des nœuds du réseau réciproque dont la position dans le réseau est du type $h\vec{a}^* + k\vec{b}^* + l\vec{c}^*$. L'autoindexation va donc déterminer trois vecteurs non-coplanaires $\vec{a}^*, \vec{b}^*, \vec{c}^*$ et de plus petit module, permettant de rendre compte de la liste de pics. Le réseau direct peut ensuite être estimé. Les relations qui existent entre \vec{a}, \vec{b} et \vec{c} imposent un choix de groupe d'espace limité.

Les données sont alors indexées image par image en affinant un certain nombre de paramètres comme la distance cristal-détecteur, la position du centre du détecteur, l'orientation du cristal, les paramètres de maille. Chaque tache est alors indexée en h, k et l. La qualité médiocre des clichés de diffraction des cristaux d'eIF3i requiert néanmoins de fixer la plupart de ces paramètres afin d'éviter une divergence du programme.

L'étape suivante des traitements consiste à intégrer les clichés de diffraction. Un fichier, produit de l'indexation, contient l'ensemble des réflexions prédites en fonction des paramètres cristallins. Les taches prédites sont ensuite alignées sur les taches enregistrées par ajustement des profils de diffraction pour chaque image. L'intensité de la tache de diffraction est mesurée en appliquant un masque autour de la tache composé de deux zones. La couronne interne délimite la surface à intégrer et la couronne externe permet d'estimer le bruit de fond à retrancher à l'intensité intégrée. Un avantage du programme Denzo comparé à la suite XDS est le traitement en mode interactif, permettant de visualiser l'intégration et notamment la cohérence du masque appliqué avec la forme des taches. L'ajustement des masques aux taches, qui est contrôlé par plusieurs paramètres, doit être optimal afin de maximiser le rapport signal/bruit. Le traitement par mosfilm est également interactif, mais la mosaïcité importante évaluée pour les cristaux d'eIF3i ($\approx 2^\circ$) limite cependant son utilisation. Seuls integrate et correct de la suite XDS et Denzo permettent de traiter les jeux de données de manière acceptable, probablement grâce à l'utilisation de profils 3D, qui étendent le masque définissant la forme des taches à la dimension « angle de rotation ».

L'étape suivante consiste à mettre à l'échelle les intensités mesurées. En effet, les images diffèrent entre elles par la dose de rayons X reçue et par la baisse du pouvoir de diffraction du cristal. Lors de la mise à l'échelle, l'intensité de chacune des réflexions indépendantes est calculée en tant que moyenne sur l'ensemble des clichés de diffraction des réflexions équivalentes à la réflexion considérée. Les programmes utilisés pour cette étape sont scala de la suite CCP4 pour des données traitées par les programmes de cette suite ou xscale pour des données traitées par XDS. Scalepack assure la mise à l'échelle de données traitées par la suite HKL.

La qualité du jeu de données est estimée par le paramètre R_{sym} et le rapport signal sur bruit (I/σ) :

$$R_{sym} = \frac{\sum_{hkl} \sum_{i=1}^N |\langle I_{hkl} \rangle - I_{hkl}^i|}{\sum_{hkl} \sum_{i=1}^N I_{hkl}^i}$$

Avec I_{hkl}^i la ième mesure de la réflexion hkl et $\langle I_{hkl} \rangle$ la valeur moyenne des N réflexions communes.

Le programme scalepack se base sur un autre critère, le χ^2 , définit par :

$$\chi^2 = \sum_{hkl} \sum_{i=1}^N \frac{(\langle I_{hkl} \rangle - I_{hkl}^i)^2}{E^2 \cdot \frac{N}{N-1}}$$

Où E est l'erreur du modèle et N est un facteur corrigeant la corrélation entre I et $\langle I \rangle$.

La pertinence de chacun des critères est toujours discutée mais le paramètre R_{sym} demeure un précieux indicateur permettant d'évaluer la qualité des données.

Enfin, les données moyennées sur les réflexions équivalentes par symétrie peuvent être réduites à l'unité asymétrique. Les amplitudes sont calculées à partir des intensités. Cette étape est plus subtile qu'une simple racine carrée car il faut tenir compte des intensités des réflexions faibles qui peuvent avoir été estimées négatives lors de la correction du bruit de fond. A cette étape, les données peuvent être mises à l'échelle absolue en calculant le facteur d'échelle et le facteur d'agitation thermique B moyen en faisant une analyse statistique de Wilson. Le logiciel truncate a été utilisé pour réaliser cette conversion. Il propose notamment des analyses statistiques des données permettant de caractériser la qualité du traitement et des données.

Dans notre cas, les premières indexations proposaient une indexation dans le groupe d'espace I222 (paramètres de mailles : 71x78x135 Å) avec une pénalité acceptable. Le traitement du meilleur jeu a permis d'obtenir un R_{sym} global de 9,6% et un R_{sym} de 35% dans la dernière tranche de résolution s'arrêtant à 2,8 Å.

4 Analyses cristallographiques préliminaires

a Contenu de la maille

Une analyse portant sur la quantité de solvant présente dans la maille permet d'évaluer le nombre de molécules dans l'unité asymétrique. Le calcul du coefficient de Matthews permet une telle détermination :

$$V_M = \frac{V}{M.Z}$$

Où V est le volume de la maille, M la masse moléculaire de la protéine et Z le nombre de molécule dans la maille. Dans notre cas, la maille I222 étant caractérisée par des angles à 90°, V est alors le produit des paramètres de mailles ($a_{I222}, b_{I222}, c_{I222}$), à savoir :

$$V = a_{I222}.b_{I222}.c_{I222} = 71 \times 78 \times 135 \text{Å}^3 = 747630 \text{Å}^3$$

La protéine a une masse de 38kDa d'où :

$$V_M = \frac{19,3}{M}$$

Le coefficient de Matthews change en fonction du taux de solvant. Pour des cristaux ayant de 30 à 75% de solvant, il varie entre 1,6 et 4,0 et, pour 50% de solvant, il avoisine 2,4. En injectant cette valeur le nombre de molécules M est de 8. La multiplicité de la maille I222 étant de 8, on en déduit qu'une molécule est présente par unité asymétrique.

b Symétries non cristallographiques, la fonction de Patterson

Lorsqu'il existe plusieurs molécules par unité asymétrique, la relation de symétrie pouvant exister entre elles n'est pas liée à la symétrie du cristal, ceci définit les symétries dites non cristallographiques (NCS). La fonction de Patterson consiste en un produit de convolution de la densité électronique avec son symétrique par rapport à l'origine. La fonction de Patterson peut ainsi s'écrire :

$$P(u, v, w) = \iiint \rho(x, y, z) \cdot \rho(u-x, v-y, w-z) dx dy dz$$

Soit, avec * le produit de convolution :

$$P(u, v, w) = \rho(x, y, z) * \rho(-x, -y, -z)$$

Or,

$$I_{hkl} \propto |F_{hkl}|^2 = F_{hkl} \cdot \overline{F_{hkl}} = F_{hkl} \cdot F_{\overline{hkl}}$$

De plus, on a vu que l'on peut écrire le facteur de structure comme la transformée de Fourier de la densité électronique. Ainsi, si on note F la transformation de Fourier :

$$I_{hkl} \propto F(\rho(x, y, z)) \cdot F(\rho(-x, -y, -z)) = F(\rho(x, y, z) * \rho(-x, -y, -z))$$

$$F^{-1}(I_{hkl}) \propto F^{-1} \cdot F(\rho(x, y, z) * \rho(-x, -y, -z)) = \rho(x, y, z) * \rho(-x, -y, -z)$$

Ainsi, $P(u, v, w) \propto F^{-1}(I_{hkl})$ et la fonction de Patterson peut-être directement calculée à partir des intensités mesurées sur les clichés de diffraction. Les cartes de Patterson contiennent alors des pics correspondant à tous les vecteurs interatomiques. Dans la fonction d'autorotation, la carte de Patterson native est superposée à elle-même après rotation. Les pics de cette fonction correspondront alors à la rotation nécessaire pour superposer les différentes molécules éventuellement présentes dans l'unité asymétrique. L'orientation de l'axe de rotation et la valeur de la rotation autour de cet axe peuvent ainsi être déterminés.

Dans un cas particulier, aucun pic de Patterson n'est observé sur la fonction d'autorotation alors qu'il existe bien une symétrie non cristallographique. Cela correspond à un axe de rotation parallèle, ou presque, à un axe cristallographique correspondant à une symétrie du même ordre (par exemple un axe d'ordre 2 parallèle à un axe cristallographique d'ordre 2, 4 ou 6). Dans ce cas, la NCS devient une translation pure. Toutefois, ces symétries non cristallographiques de translation pure peuvent être repérées car elles vont provoquer un pic dans la fonction de Patterson.

La présence présumée d'une seule molécule par unité asymétrique ne nous destine pas *a priori* à réaliser ces analyses. Toutefois, à notre grande surprise, nous avons constaté la présence d'un pic intense dans la fonction de Patterson sur la section $z=1/2$. En calculant la fonction à différentes résolutions, nous avons pu déduire une translation pure d'une demi-maille selon l'axe a et d'une

demi-maille suivant \vec{c} (le pic relevé ayant pour coordonnées (x=1/2, y=0, z=1/2)). Ceci nous a donc fait nous interroger sur la détermination du groupe d'espace puisque la présence de deux molécules par unité asymétrique est incompatible avec les paramètres de maille du groupe d'espace I222. De plus, la redissolution des cristaux d'eIF3i montre que la protéine présente est intacte. Une analyse plus détaillée de la symétrie du jeu de données réalisée avec le programme pointless, récemment implémenté à la suite CCP4, suggère la présence d'un seul vrai axe de symétrie d'ordre 2, et non de trois axes comme en I222. La symétrie réelle serait donc monoclinique et non orthorhombique.

c Détermination du groupe d'espace correct et contrôle des données

Les informations précédentes nous indiquent que notre maille ne comporte qu'un seul axe d'ordre 2. Les données peuvent effectivement être indexées dans le groupe C2. Les paramètres de mailles deviennent alors :

$$(a ; b ; c ; \alpha ; \beta ; \gamma) = (156,7 ; 71,6 ; 77,9 ; 90,0 ; 119,7 ; 90,0)$$

$$\text{Où : } a = \|\vec{a}\| ; b = \|\vec{b}\| ; c = \|\vec{c}\| ; \alpha = (\vec{b}, \vec{c}) ; \beta = (\vec{a}, \vec{c}) ; \gamma = (\vec{a}, \vec{b})$$

Alors que dans ce type de réseau, les paramètres de maille a, b, c et β sont quelconques, la maille d'eIF3i présente les particularités : $a \approx 2c$ et $\beta \approx 120^\circ$.

Intéressons-nous au contenu de la maille :

$$V_M = \frac{a.b.c.\cos(\beta - 90^\circ)}{M.Z} = \frac{756665 \text{ \AA}^3}{M.Z} = \frac{19,5}{Z}$$

Ainsi, le volume de la maille est identique à celle du traitement dans le groupe d'espace I222. 8 molécules seraient donc présentes. Puisque la maille C2 assure une multiplicité de 4, on en déduit qu'il y a 2 molécules par unité asymétrique. De façon tout à fait cohérente, la NCS de translation pure est toujours détectable et correspond dans ce système à une translation d'une demi-maille suivant l'axe a . Les deux molécules présentes dans l'unité asymétriques sont donc liées par une translation d'un vecteur (1/2,0,0).

d Emulation de la maille I222 par la maille C2

Afin de montrer la concordance entre les deux systèmes, définissons un repère orthonormé $(\vec{u}, \vec{v}, \vec{w})$ de l'espace réel tel que :

$$\vec{a}_{C2} = \|\vec{a}_{C2}\|.\vec{u} ; \vec{b}_{C2} = \|\vec{b}_{C2}\|.\vec{v} \text{ et } \vec{c}_{C2} = -\|\vec{c}_{C2}\|.\sin(30^\circ).\vec{u} + \|\vec{c}_{C2}\|.\cos(30^\circ).\vec{w}$$

Soit M (x,y,z) un nœud du réseau et O l'origine de la maille :

$$\vec{OM} = x.\vec{a}_{C2} + y.\vec{b}_{C2} + z.\vec{c}_{C2}$$

Or les paramètres de la maille permettent à la diagonale de \vec{a}_{C2} et \vec{c}_{C2} d'être normale au vecteur \vec{c}_{C2} :

$$\vec{c}_{C2} \cdot (\vec{a}_{C2} + \vec{c}_{C2}) = 0$$

Et on peut écrire :

$$\overrightarrow{OM} = x \cdot (\overrightarrow{a_{C2}} + \overrightarrow{c_{C2}}) + y \cdot \overrightarrow{b_{C2}} + (z - x) \cdot \overrightarrow{c_{C2}}$$

Or, en considérant que $\|\overrightarrow{a_{C2}}\| \approx 2 \cdot \|\overrightarrow{c_{C2}}\|$, on obtient :

$$(\overrightarrow{a_{C2}} + \overrightarrow{c_{C2}}) \approx \frac{3 \cdot \|\overrightarrow{c_{C2}}\|}{2} \cdot \overrightarrow{u} + \frac{\sqrt{3} \cdot \|\overrightarrow{c_{C2}}\|}{2} \cdot \overrightarrow{w}$$

En remarquant : $\|\overrightarrow{a_{C2}} + \overrightarrow{c_{C2}}\| = \sqrt{3} \cdot \|\overrightarrow{c_{C2}}\| \approx \|\overrightarrow{c_{I222}}\|$; $\|\overrightarrow{b_{C2}}\| \approx \|\overrightarrow{b_{I222}}\|$; $\|\overrightarrow{c_{C2}}\| \approx \|\overrightarrow{a_{I222}}\|$;
on peut alors écrire :

$$\overrightarrow{OM} = (z - x) \cdot \overrightarrow{a_{I222}} + y \cdot \overrightarrow{b_{I222}} + x \cdot \overrightarrow{c_{I222}}$$

$$\text{avec : } \overrightarrow{c_{I222}} = (\overrightarrow{a_{C2}} + \overrightarrow{c_{C2}}) \quad ; \quad \overrightarrow{b_{I222}} = \overrightarrow{b_{C2}} \quad ; \quad \overrightarrow{a_{I222}} = \overrightarrow{c_{C2}}$$

Ainsi, l'indexation dans une maille I222 correspond à un changement de repère dont les axes sont liés à ceux de la maille C2 par la relation précédente et qui transforme les coordonnées d'un nœud M (x,y,z) en (z-x,y,x) (**Figure 3.17**).

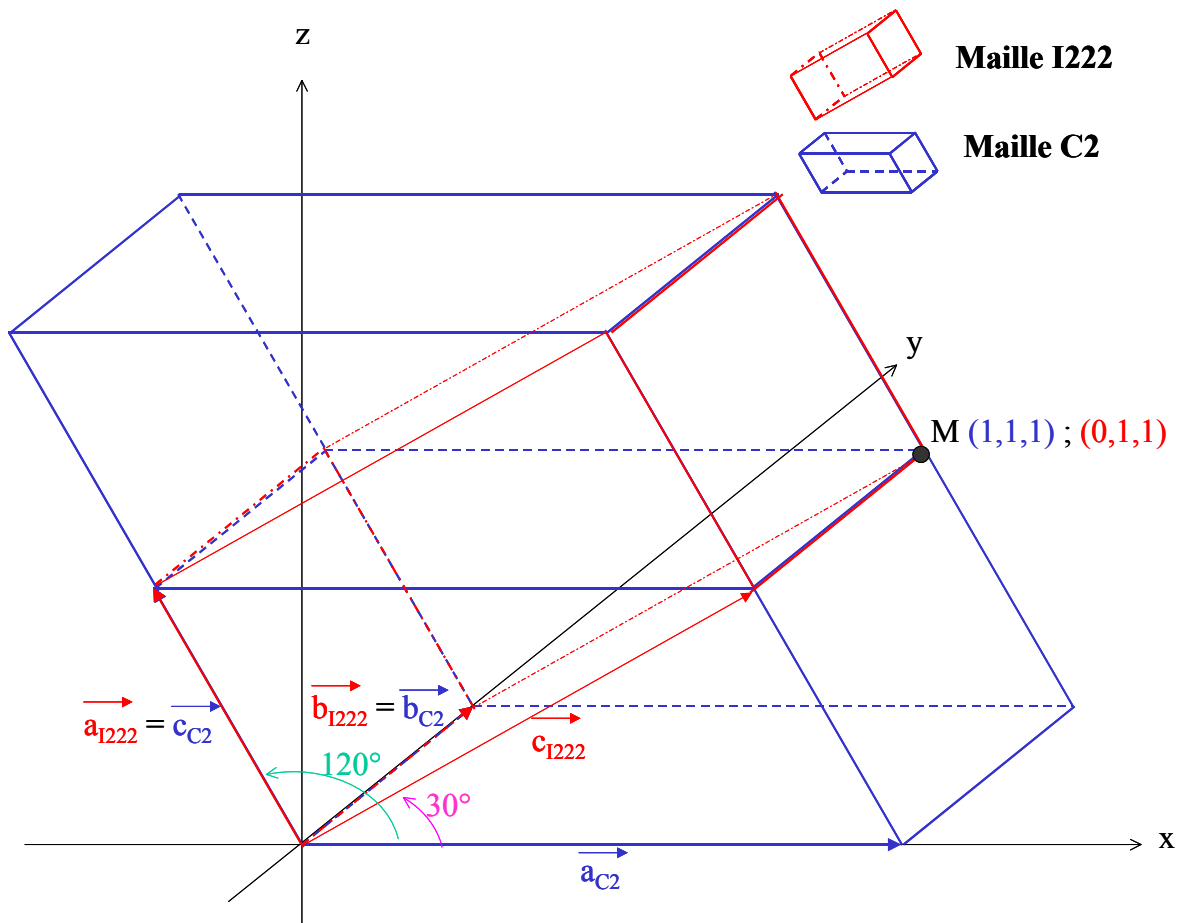


Figure 3.17 : Représentation des mailles cristallographiques I222 et C2.

La maille I222 telle qu'elle est décrite ne présente qu'un unique axe d'ordre 2 commun avec la maille C2 (axe $\overline{c_{C2}}$ et $\overline{a_{I222}}$). Nous en déduisons qu'un autre phénomène est à l'origine de la symétrie apparente I222. Or, lorsque l'indexation est réalisée en C2, l'analyse de la qualité des données par le logiciel truncate indique la présence d'un jumelage mérohédral du cristal.

e Jumelage : origine et conséquence

Le jumelage correspond à un défaut de croissance d'un cristal au sein duquel plusieurs mailles identiques ayant des orientations différentes coexistent. Dans notre cas, les positions des taches de diffraction correspondant aux deux réseaux cristallins se superposent exactement, on parle alors de jumelage mérohédral. Les différentes mailles sont alors liées par une symétrie qui n'appartient pas à celle de chacune des mailles, ce qui va donc générer des symétries supplémentaires. La proportion de la maille 2, jumelle, par rapport à la maille 1, native, est la fraction de jumelage (toujours inférieure à $\frac{1}{2}$). Lorsque la fraction de jumelage approche de la valeur $\frac{1}{2}$, le jumelage est dit parfait. L'intensité observée résulte de la somme des intensités correspondant à chacun des deux réseaux :

$$I_{obs}(hkl) = (1 - \alpha) \cdot I(h_1k_1l_1) + \alpha \cdot I(h_2k_2l_2)$$

Plusieurs logiciels calculent des statistiques sur les réflexions observées afin d'évaluer la fraction de jumelage et de déterminer l'opération de symétrie, ou loi de jumelage, liant les mailles jumelles. La loi de jumelage est une opération réalisée dans l'espace réciproque. Nous avons utilisé principalement le programme phenix.xtriage pour déterminer les paramètres de jumelage. Notamment, le logiciel trace un diagramme de Britton qui reporte le nombre d'intensités négatives du jeu de données déjumelé en fonction de la fraction de jumelage. Si α est supérieur à sa valeur réelle alors trop d'intensités des réflexions de la maille jumelée seront retranchées à celles de la maille native, induisant l'apparition d'intensités négatives. Le tracé reste plat jusqu'à une valeur de la fraction de jumelage où le nombre d'intensités négatives augmente rapidement. L'intersection de cette asymptote avec l'axe où est reporté α apparaît comme un moyen possible pour évaluer sa valeur réelle. Dans notre cas, la fraction de jumelage est évaluée à 0,2 pour tous les jeux.

La loi de jumelage déterminée par le programme est $(-h-2l, -k, l)$. Afin de comprendre ce qu'elle représente, il faut caractériser la transformation dans l'espace réciproque. Soit un point M (h, k, l) un nœud du réseau réciproque et M' son jumeau. Alors :

$$\overline{OM} = h \cdot \overline{a^*} + k \cdot \overline{b^*} + l \cdot \overline{c^*} \quad (1)$$

$$\overline{OM'} = (-h-2l) \cdot \overline{a^*} - k \cdot \overline{b^*} + l \cdot \overline{c^*}$$

Soit :

$$\overline{OM'} = -h \cdot \overline{a^*} - k \cdot \overline{b^*} + l \cdot (\overline{c^*} - 2\overline{a^*})$$

Ainsi, le point M' est obtenu par une rotation de 180° autour de l'axe $\overline{c^*}$ où $\overline{c'^*} = \overline{c^*} - 2\overline{a^*}$ dans le repère $(\overline{a^*}, \overline{b^*}, \overline{c'^*})$.

Soit (H,K,L) les coordonnées de M dans ce repère :

$$\begin{aligned}\overline{OM} &= H.\overline{a^*} + K.\overline{b^*} + L.(\overline{c^*} - 2\overline{a^*}) \\ \overline{OM} &= (H - 2L).\overline{a^*} + K.\overline{b^*} + L.\overline{c^*} \quad (2)\end{aligned}$$

Par identification entre les équations (1) et (2), on obtient :

$$\begin{cases} H = h + 2l \\ K = k \\ L = l \end{cases} \quad [1]$$

Ainsi, la transformation consiste à effectuer un changement de repère puis une rotation autour de $\overline{c^{i^*}}$:

$$(h, k, l) \xrightarrow{(\overline{a^*}, \overline{b^*}, \overline{c^{i^*}})} (h + 2l, k, l) \xrightarrow{\text{rotation}} (-h - 2l, -k, l)$$

Afin de reporter ceci dans l'espace direct, il faut caractériser le vecteur $\overline{c^{i^*}}$. Plaçons-nous dans le repère orthonormé $(\overline{u}, \overline{v}, \overline{w})$ tel que :

$$\overline{b^*} = \|\overline{b^*}\|.\overline{v^*} \text{ et } \overline{c^*} = \|\overline{c^*}\|.\overline{w^*}$$

Alors, d'après les paramètres de maille, $(\overline{a^*}, \overline{u^*}) = 60^\circ$ et :

$$\overline{a^*} = \frac{\|\overline{a^*}\|}{2}.\overline{w^*} + \frac{\sqrt{3}\|\overline{a^*}\|}{2}.\overline{u^*} \quad (3)$$

Par définition de l'espace réciproque, on a :

$$\|\overline{a^*}\| = \frac{\|\overline{b}\|.\|\overline{c}\|}{V} = \frac{2}{\sqrt{3}.\|\overline{a}\|} \quad (4) \text{ et } \|\overline{c^*}\| = \frac{\|\overline{a}\|.\|\overline{b}\|}{V} = \frac{2}{\sqrt{3}.\|\overline{c}\|} \approx \frac{4}{\sqrt{3}.\|\overline{a}\|} \quad (5)$$

En combinant les équations (3), (4) et (5) on peut écrire le vecteur $\overline{c^{i^*}}$ dans le repère orthonormé :

$$\|\overline{c^{i^*}}\| = \|\overline{c^*}\|.\left(\cos(-60^\circ).\|\overline{w^*}\| + \sin(-60^\circ).\|\overline{u^*}\|\right)$$

Le vecteur $\overline{c^{i^*}}$ correspond ainsi au vecteur $\overline{c^*}$ tourné de -60° dans le plan $(0, \overline{u^*}, \overline{w^*})$. Or une rotation dans l'espace réciproque étant une rotation dans l'espace réel, le vecteur $\overline{a^i}$ de l'espace correspondant à $\overline{c^{i^*}}$ est obtenu par la même rotation dans le plan $(0, \overline{u}, \overline{w})$ du vecteur \overline{a} . La maille jumelle peut ainsi être obtenue en appliquant une rotation à la maille native de 180° par rapport à $\overline{a^i}$. C'est précisément cet axe d'ordre 2 que crée le jumelage, qui est à l'origine de la symétrie apparente I222.

De plus, il peut être remarqué que tout nœud M' du réseau jumeau s'écrit sous la forme :

$$\overrightarrow{OM'} = h.\overrightarrow{(-a^*)} + k.\overrightarrow{(-b^*)} + l.\overrightarrow{c^*}$$

Ainsi le repère du réseau réciproque jumeau est $(\overrightarrow{-a^*}, \overrightarrow{-b^*}, \overrightarrow{c^*})$. Il est donc possible de lui faire correspondre le repère du réseau jumeau dans l'espace direct : $(\overrightarrow{-a}, \overrightarrow{-b}, \overrightarrow{c})$ et d'y dessiner ainsi la maille jumelle (**Figure 3.18**).

En outre, le changement de repère identifié par le système d'équation [1] correspond à une autre indexation possible des données. Afin d'avoir des jeux de données homogènes, les données ont toutes été indexées dans le même système, ce qui a pu être vérifié à l'aide du logiciel pointless.

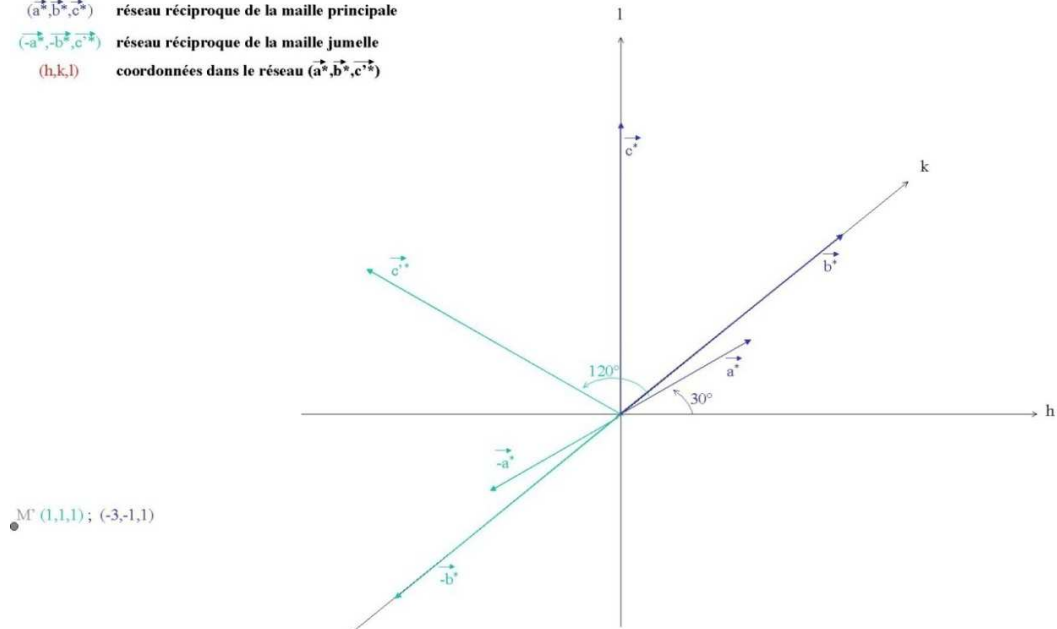
Le travail avec des données jumelées rend difficile l'utilisation de la méthode MIR pour phaser les données en raison de la forte imprécision sur la valeur de la fraction de jumelage. Très peu de structures ont pu être déterminées à partir de données jumelées par cette méthode de phasage ([Cheah et al., 1994](#); [Declercq and Evrard, 2001](#); [Fisher and Sweet, 1980](#); [Igarashi et al., 1997](#); [Rees and Lipscomb, 1980](#); [Valegard et al., 1998](#)).

La méthode de remplacement moléculaire reste possible. En effet, un traitement par cette technique devrait, dans le meilleur des cas, fournir les 2 solutions correspondant à chacune des mailles.

Espace réciproque

● $M(1,1,1); (3,1,1)$

- $(\vec{a}^*, \vec{b}^*, \vec{c}^*)$ réseau réciproque de la maille principale
- $(-\vec{a}^*, -\vec{b}^*, \vec{c}^*)$ réseau réciproque de la maille jumelle
- (h, k, l) coordonnées dans le réseau $(\vec{a}^*, \vec{b}^*, \vec{c}^*)$



Espace réel

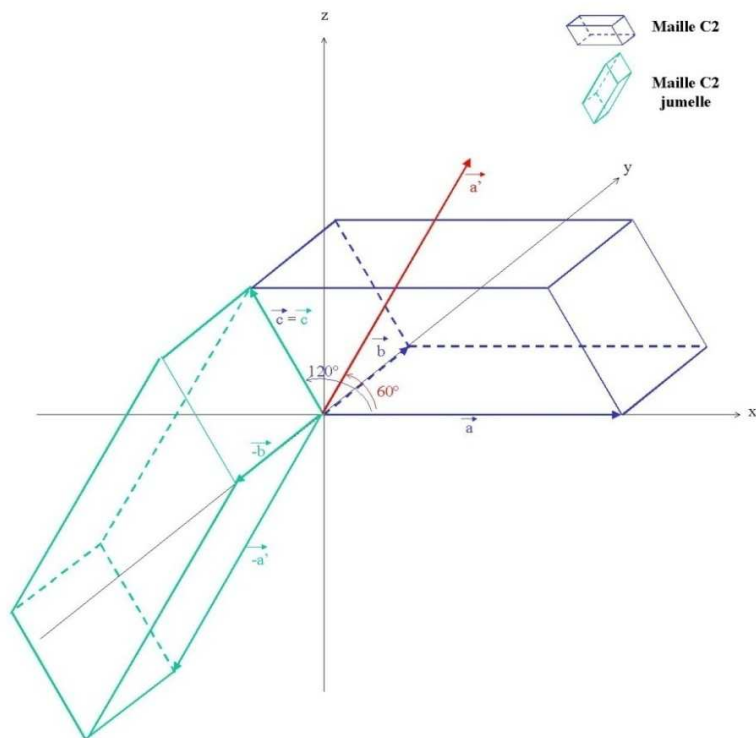


Figure 3.18 : Représentation de la maille cristallographique C2 et de sa maille jumelle.

5 Phasage des données

a Phasage par remplacement moléculaire

Principe

Cette méthode permet de déterminer les phases en positionnant un modèle homologue à la protéine cristallisée. En effet, il peut être généré à partir d'une structure tridimensionnelle un diagramme complet de diffraction incluant amplitudes et phases. Ainsi, si le diagramme de diffraction expérimental et le diagramme de diffraction calculé sont comparables, il peut être déterminé un jeu de phases initiales. Pour que cette superposition soit réalisable, il faut cependant que les deux structures occupent une position identique dans la maille cristallographique. Les programmes de remplacement moléculaires visent ainsi à placer la molécule modèle dans la maille de la structure à déterminer. Le programme calcule 6 paramètres (3 angles d'orientation et 3 vecteurs de position) en deux étapes successives de détermination de l'orientation (rotation) puis de la position (translation).

La fonction de rotation évalue l'accord entre les fonctions de Patterson du modèle et de la structure à déterminer, en fonction de leurs orientations relatives :

$$R(C) = \iiint_{\mathcal{V}} P_{nat}(\vec{r}) \cdot U(\vec{r}) \cdot P_{mod}([C]\vec{r}) \cdot d\mathcal{V}$$

Où :

[C] est la matrice de rotation

$P_{nat}(\vec{r})$ est la valeur de la fonction de Patterson de la molécule de la structure à résoudre en \vec{r}

$P_{mod}([C]\vec{r})$ est la valeur de la fonction de Patterson du modèle en $[C]\vec{r}$

$U(\vec{r})$ est une fonction qui vaut 1 à l'intérieur du rayon d'intégration et 0 à l'extérieur

Le volume d'intégration permet de ne sélectionner que les vecteurs interatomiques au sein d'une molécule, éliminant ainsi les vecteurs intermoléculaires. La fonction de rotation sera ainsi maximale lorsque les deux fonctions de Patterson sont superposées par l'opération C. Ce calcul est réalisé par l'algorithme de fonction de rotation rapide où la fonction de Patterson est développée en une somme d'harmoniques sphériques permettant le calcul par la méthode de la transformée de Fourier rapide. Les logiciels de remplacement moléculaire identifient ainsi une liste de rotations correspondant aux maxima de la fonction R(C) et attribuent un score à chacune de ces rotations. Généralement, la solution correcte présente un pic de la fonction de rotation qui se détache des autres rotations trouvées.

La seconde étape consiste à déterminer la position de la structure modèle dans la maille. La fonction de translation prend la forme :

$$T(\vec{t}) = \int_v [P_{nat}(u) - P_{nat,intra}(u)] \cdot P_{mod,rot}(u, \vec{t}) \cdot dv$$

Avec :

P_{nat} est la fonction de Patterson de la molécule de la structure à déterminer

$P_{nat,intra}$ correspond aux vecteurs intramoléculaires

$P_{mod,rot}$ est la fonction de Patterson calculée grâce au modèle auquel a été appliqué la rotation C

La fonction $T(\vec{t})$ est alors maximale pour toute valeur de \vec{t} correspondant à un vecteur intermoléculaire, entre molécules symétriques par exemple. La fonction $T(\vec{t})$ est évaluée par une simple transformée de Fourier.

La qualité des solutions ainsi déterminées est estimée par la valeur des coefficients de corrélation C sur les intensités diffractées observées et calculées à partir du modèle dans sa nouvelle orientation et position. La valeur du facteur R et l'empilement cristallin constituent également des critères de qualité des solutions.

$$C = \frac{\sum_h \sum_k \sum_l [|F_{obs}|^2 - \overline{|F_{obs}|^2}] \times [|F_{calc}|^2 - \overline{|F_{calc}|^2}]}{\sqrt{\sum_h \sum_k \sum_l [|F_{obs}|^2 - \overline{|F_{obs}|^2}]^2 \times \sum_h \sum_k \sum_l [|F_{calc}|^2 - \overline{|F_{calc}|^2}]^2}}$$

$$R = \frac{\sum_h \sum_k \sum_l \| |F_{obs}| - k |F_{calc}| \|}{\sqrt{\sum_h \sum_k \sum_l |F_{obs}|}}$$

Où k est le facteur de mise à l'échelle entre F_{obs} et F_{calc} .

Il est à noter que les programmes les plus récents, comme phaser, utilisent des méthodes de classement des solutions plus fines, basées sur des critères de maximum de vraisemblance ([Storoni et al., 2004](#)).

Le remplacement moléculaire pour eIF3i

Les alignements des séquences d'eIF3i font apparaître l'homologie de cette protéine à toute une famille de protéines (25% d'identité de séquence avec WDR5, le meilleur score) : les hélices β (β -propeller). Ce motif β -propeller se retrouve chez de nombreuses protéines puisqu'il s'avère être le second motif le plus rencontré après le repliement de Rossman ([Smith et al., 1999](#)). Ce repliement n'est formé que de feuillets β . Les structures présentent un axe de symétrie qui lie des motifs en lame (blade) composés de 4 brins antiparallèles (**Figure 3.19**). Le nombre de motifs lame varie de 4 à 8 et est le plus souvent de 7. Les structures à 7 lames présentent d'ailleurs moins de variabilité comparées aux autres types d'hélices β . Alors que seuls cinq motifs lames avaient été prédits dans eIF3i lors de son identification, les études suivantes font état de 7 motifs dans la structure du facteur humain alors que les alignements de séquences sont plutôt cohérents avec 7 répétitions (**Figure 3.20**).

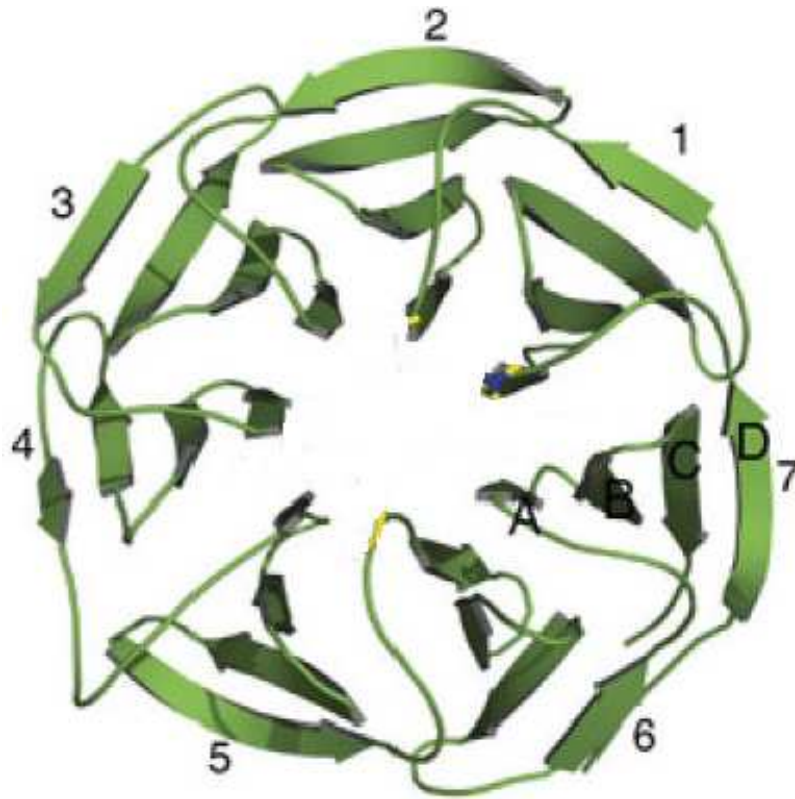


Figure 3.19 : Structure tridimensionnelle de WDR5 d'après ([Schuetz et al., 2006](#)). Les 7 motifs lames caractéristiques des β -propeller sont numérotés de 1 à 7 et les brins du septième motif sont nommés de A à D, du brin interne au brin externe.

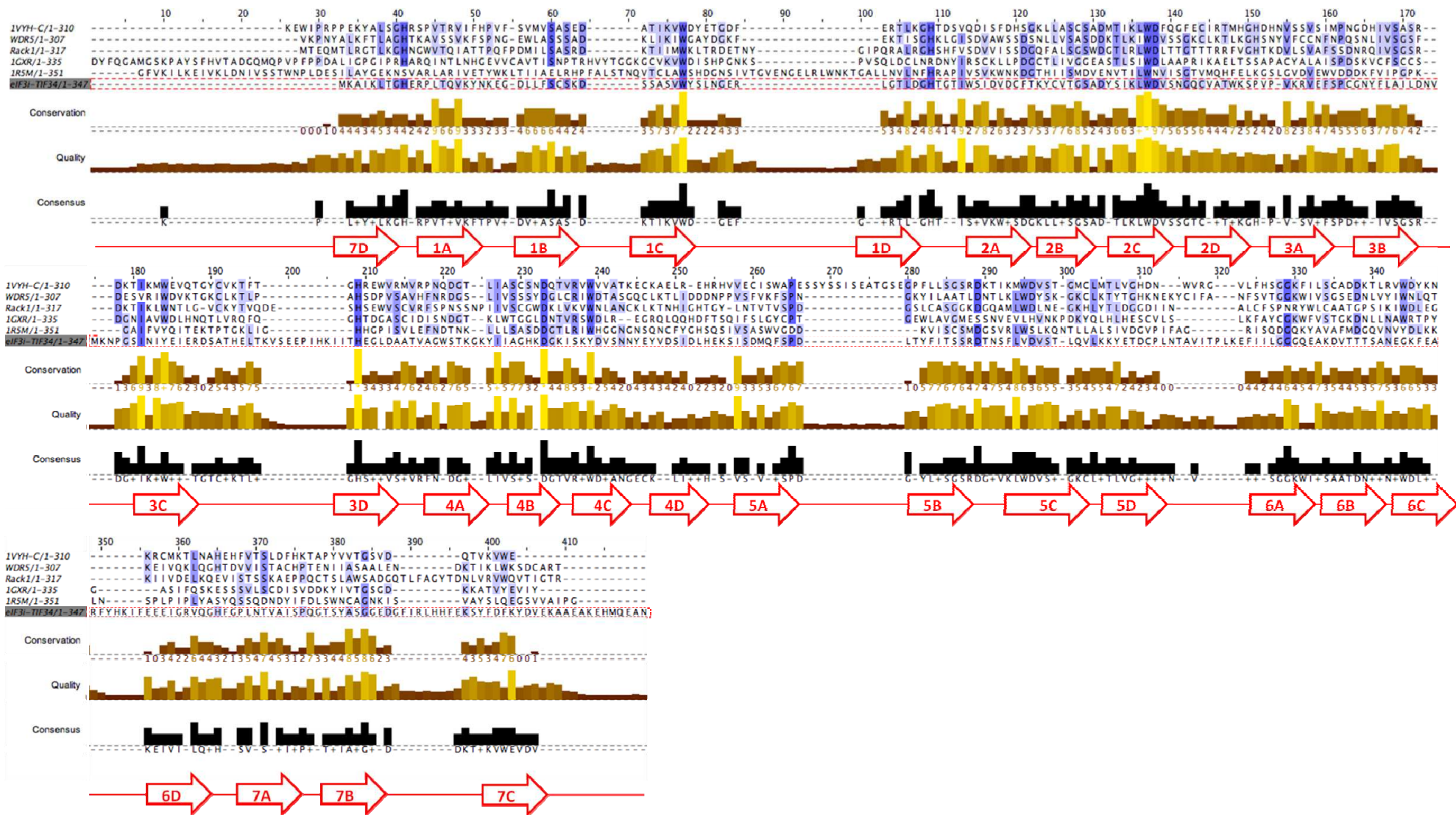


Figure 3.20 : Alignement de séquences de β -propeller et d'eIF3i réalisé avec Clustal W (Chenna et al., 2003) et modifié en tenant compte de la superposition des structures 3D de WDR5, 1VYH, 1R5M et 1GXR. La conservation des résidus est représentée par une échelle de bleu, le bleu foncé représentant une forte conservation. La séquence d'eIF3i de levure est encadrée en rouge. Les protéines sont nommées en fonction de leurs noms (eIF3i, WDR5, Rack1) ou de leur identifiant pdb (1VYH, 1R5M). La structure secondaire consensus des structures résolues est reportée en rouge en dessous de l'alignement. La nomenclature utilisée pour les brins correspond à celle des β -propeller (Figure 5.3)

Différentes approches ont été mises en œuvre :

1. Les différentes structures à 5, 6 et 7 motifs lames ont été testées dans les trois programmes de remplacement moléculaire à notre disposition : phaser, molrep et Amore de la suite CCP4.
2. Les structures de représentants des différentes sous-familles ont été superposées afin de créer pour chaque sous-famille un cœur commun minimal, utilisé comme modèle pour la recherche dans les différents programmes.
3. Une autre approche a consisté à ne rechercher qu'un seul motif lame. Ceci a permis d'obtenir une solution mais la recherche du motif suivant n'a jamais abouti.
4. Les suites de programmes CaspR (utilisant Amore) et MrBUMP (basé sur molrep et phaser) permettant une automatisation de la procédure ont ensuite été testées. CaspR nécessite d'indiquer des structures modèles à partir desquelles des structures moyennées, déterminées à l'aide des programmes T-coffee et Modeller, sont essayées dans le remplacement moléculaire. Bien qu'ayant testé les différentes sous-familles, les solutions présentant les meilleurs facteurs C et R ne satisfaisaient pas le critère d'empilement cristallin. La seconde suite se charge de sélectionner les modèles à utiliser grâce à des alignements de séquence, et effectue des troncatures de ces modèles déduites des alignements. Là encore, toutes les tentatives ont échoué.
5. Le traitement de la NCS est pris en charge par les logiciels Amore et Molrep qui, après avoir déterminé la fonction de rotation, cherchent la translation en partant d'un modèle comprenant les deux molécules reliées par l'opération de symétrie non cristallographique. Le logiciel phaser fournissant souvent de meilleurs résultats, une démarche similaire a été tentée. Un fichier utilisé pour la recherche de la translation a été créé en y insérant les coordonnées des modèles correspondant aux meilleures fonctions de rotation et les coordonnées de leurs images par la NCS. Cette manipulation n'a cependant pas permis d'obtenir de solution.
6. L'intervention sur les différents paramètres de recherche modifiables dans les programmes (résolution, forme des fonctions de recherche, critère de sélection,...) a été testée dans les différentes situations précédentes.
7. Le jumelage au sein des cristaux n'est *a priori* pas un facteur rédhibitoire pour le phasage par remplacement moléculaire puisque les fractions des réflexions provenant de la maille la plus présente devraient fournir le signal le plus fort et donc permettre une résolution de l'emplacement de la structure modèle. Les difficultés rencontrées doivent certainement être un effet conjugué du jumelage, de la NCS de translation pure, de la symétrie des structures modèles (qui doit limiter la détermination de la fonction de rotation) et de l'absence de modèle fiable. L'influence du premier problème a été testée en tentant de « déjumeler » les données à l'aide du logiciel detwin de la suite CCP4 qui génère un jeu de données de la maille native en soustrayant les intensités de la maille jumelle, pondérées par la fraction de jumelage. Ceci est en général peu conseillé car il conduit à une atténuation des intensités, intensifiant encore l'effet du jumelage qui a tendance à moyenniser les intensités des jeux de données de la maille native et de la maille jumelle. Aucune amélioration du signal n'a effectivement pu être observée.

Malgré les différentes tentatives, aucun traitement n'a fourni de solution avec des critères acceptables.

b Phasage par MIR

Principe

Le remplacement isomorphe multiple (MIR) consiste à ajouter un atome lourd dans le cristal, sans en altérer la maille cristalline. Dans ce cas, le facteur de structure enregistré pour le cristal dérivé est la somme vectorielle dans l'espace des nombres complexes du facteur de structure du cristal natif et de l'atome lourd.

$$F_{PH} = F_P + F_H$$

Où :

F_{PH} est le facteur de structure du cristal dérivé

F_P représente le facteur de structure du cristal natif

F_H est le facteur de structure de l'atome lourd

Les amplitudes des différents facteurs de structures sont atteintes indirectement par la mesure des intensités. Le paramètre de phase est représenté dans cet espace par l'angle entre le vecteur F et l'axe des nombres réels. Seul le vecteur F_H peut être déterminé grâce au positionnement dans la maille de l'atome lourd par les méthodes de Patterson. Ainsi après avoir tracé le cercle de rayon $|F_{PH}|$, un cercle identique est tracé à partir du vecteur $-F_H$. Les deux points d'intersections de ces cercles désignent deux valeurs possibles du vecteur F_P . La combinaison avec l'enregistrement d'un nouveau dérivé permettra de lever l'indétermination sur F_P (**Figure 3.21**).

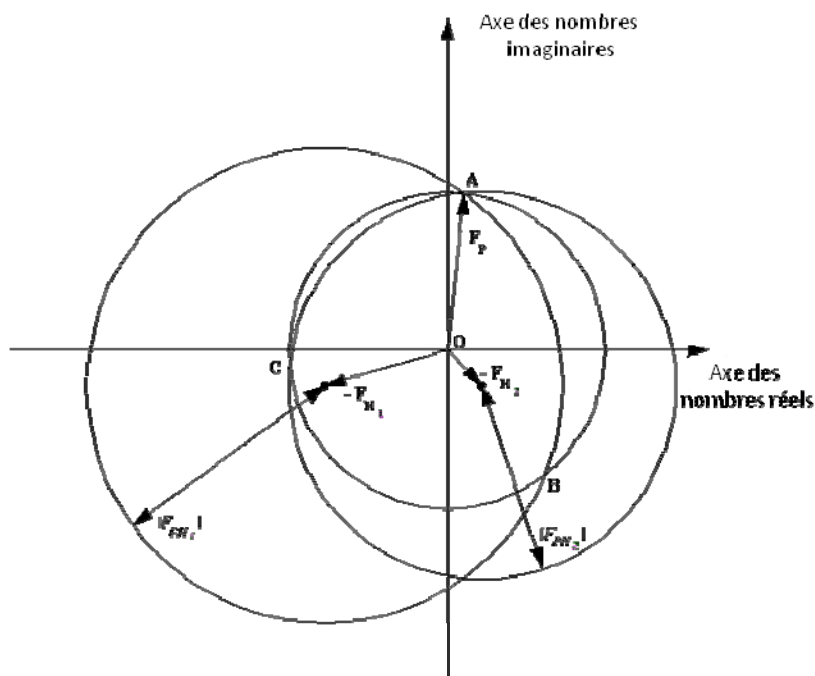


Figure 3.21 : Visualisation de la détermination des phases par MIR. Un premier cercle de diamètre $|F_P|$ est tracé. A partir d'un premier dérivé, le cercle de diamètre $|F_{PH1}|$ à partir de $-F_{H1}$ intercepte le premier cercle à deux endroits indiquant deux valeurs de phase possibles. Le tracé d'un troisième cercle à partir d'un deuxième dérivé lourd indique l'unique point de concours des trois cercles (A), fournissant sans ambiguïté la phase.

Les dérivés sont obtenus généralement par trempage des cristaux dans des solutions contenant l'atome lourd. Les composés lourds diffusent dans le cristal à travers les canaux de solvant et peuvent se fixer en des sites précis dépendants de l'environnement chimique. Par exemple, les composés à base de mercure réagissent plus spécifiquement avec les résidus cystéines accessibles. Le critère permettant de caractériser un « bon » dérivé, R_{iso} , évalue l'isomorphisme entre les données du cristal natif et les données du cristal dérivé.

$$R_{iso} = \frac{\langle |F_{PH}| - |F_P| \rangle}{\langle |F_P| \rangle} = \sqrt{\frac{N}{2}} \times \frac{^0f}{\langle |F_P| \rangle}$$

Avec :

N le nombre d'atomes lourds

0f le facteur de diffusion de l'atome lourd

$\langle |F_P| \rangle$ le facteur de diffusion moyen de la protéine

La position des atomes lourds peut être simplement déterminée par des différences entre les fonctions de Patterson du cristal natif et du cristal dérivé. Le tracé des différentes sections de Harker de ces fonctions doit alors mettre en évidence les vecteurs entre les atomes lourds.

La partie délicate du phasage consiste alors à affiner la position des atomes lourds dans la maille. Le modèle d'atomes lourds est complété en introduisant les sites d'atomes lourds mineurs trouvés par différence de Fourier en utilisant les phases calculées à partir des positions d'atomes lourds déjà identifiés. La fonction différence de Fourier est dans ce cas une synthèse de Fourier calculée avec les différences de modules ($\langle |F_{PH}| - |F_P| \rangle$) et les phases calculées, φ_{calc} , à l'aide le dérivé. Les programmes mlphare ([Otwinoski, 1991](#)) et, de manière plus fine, sharp ([La Fortelle and Bricogne, 1997](#)) utilisent les méthodes de maximum vraisemblance pour modéliser les erreurs. Avant l'extraction des phases, le modèle d'atomes lourds doit être affiné avec précision ainsi que le défaut d'isomorphisme. Le système de maximum de vraisemblance où les erreurs sont supposées avoir une répartition gaussienne permet de mieux modéliser les erreurs en associant une distribution de probabilité à toutes les grandeurs à affiner. Le résultat de l'affinement n'est donc pas une valeur des phases mais, pour chacune d'elles, une distribution de probabilité tenant compte de toutes les informations. Dans le programme sharp, la construction de la fonction de vraisemblance introduit une hiérarchie entre les paramètres, définie par l'utilisateur. Le défaut d'isomorphisme est dans ce cas un paramètre affiné, contrairement à mlphare où il est évalué à chaque cycle. Le programme sharp constitue ainsi un outil puissant pour les données de faibles intensités.

A l'issue de ce traitement, les phases de chaque facteur de structure sont estimées et se répartissent suivant une fonction de probabilité $P(\varphi)$. La phase la meilleure, φ_{best} , est la moyenne pondérée par la probabilité $P(\varphi)$ des phases possibles et représente alors le centroïde de la distribution de probabilité.

$$F_{best} = \frac{\sum_{\varphi} P(\varphi) \cdot F(\varphi)}{\sum_{\varphi} P(\varphi)}$$

$$F_{hkl,best} = |F_{hkl}| \cdot m_{hkl} \cdot e^{i\varphi_{best}} \text{ avec } m = \frac{|F_{hkl,best}|}{|F_{hkl}|}$$

Le facteur m reflète ainsi l'étalement de la distribution de probabilité et est appelée figure de mérite (Figure Of Merit : FOM). Une carte de densité électronique peut être calculée avec les $|F_p|$ et les phases expérimentales φ_{best} .

Obtention de dérivés et phasage

De nombreuses solutions de cristallisation complétées avec un composé lourd compatible avec le pH de la solution ont été essayées. Les composés qui ont été utilisés et les données éventuellement collectées sont reportés dans le **tableau 3.5**. Les tracés des sections Harker des jeux traités ne mettent cependant pas en évidence de pics sortant du bruit. Il est à noter que le jumelage des cristaux est souvent un frein au phasage par cette technique (**voir p. 132**). Nous avons tenté de travailler à partir de données « déjumelées », mais cela n'a pas permis d'obtenir plus d'informations. Les dérivés au mercure, et notamment le pcmb, ont largement été testés. Cependant, un trempage de 20 heures des cristaux dans une solution à 2mM fait éclater les cristaux. Un trempage à une concentration de 0,5mM pendant 20 heures n'a pas altéré la maille cristalline, mais l'analyse des sections de Harker des jeux enregistrés pour les cristaux restés intacts ne fait pas apparaître de sites clairement identifiables. Quelques tentatives de phasage avec les pics de Patterson les plus intenses n'ont pas permis d'aboutir.

La protéine eIF3i contient 6 cystéines. Il existe donc de nombreux sites de liaisons putatifs pour le mercure. Si les atomes de mercure se lient à tous les résidus cystéines, il est possible qu'à faible concentration en mercure (0,5mM), le signal de liaison ne soit pas visible car réparti sur les 6 atomes et qu'à forte concentration, les cristaux s'altèrent en raison d'un taux de dérivation trop important. Nous avons donc essayé de muter ces cystéines afin de ne laisser que deux sites de liaison pour le mercure. Le choix des cystéines à muter dérive d'une superposition de la séquence d'eIF3i sur les structures homologues résolues. Quatre cystéines possiblement accessibles au solvant (C62, C67, C104 et C238) ont ainsi été sélectionnées. Les deux premières étant proches l'une de l'autre, il a été décidé de n'en laisser qu'une seule sur les deux, les autres cystéines ont été changées en alanine ou en serine (**voir Matériels et méthodes**). La méthode QuikChange® Multi-Site Directed Mutagenesis (Stratagene) permet de réaliser des mutations multiples en une seule étape. Deux mutants d'eIF3i contenant comme seules cystéines les résidus 62, 104 et 238, d'une part, ou 67, 104 et 238, d'autre part, ont ainsi été obtenus. Seul le mutant comportant la cystéine 62 a cependant pu être purifié de manière homogène. Les essais de cristallisation avec la condition de cristallisation adaptée à la protéine native ainsi que celles des matrices disponibles n'ont cependant pas permis d'obtenir de cristaux.

Une autre piste que nous développons consiste à tenter de cristalliser la protéine dans la condition de cristallisation additionnée d'une très faible quantité d'atome lourd. Nous devrions bientôt savoir si cette piste permettra d'obtenir des cristaux.

Atome	Composé	Concentration	Temps de trempage	Résolution	Rsym (global)	Rsym (Réso. Max.)
Mercure	pcmbs	2mM	3h15	4,1	17,40%	33,40%
		2mM	20h		pas de diffraction	
		0,5mM	20h	2,8	8,30%	37,70%
	HgBr ₂	1mM	20h		pas de diffraction	
		2mM	20h		pas de diffraction	
	HgCl ₂	0,5mM	7h		pas de diffraction	
		2mM	20h		pas de diffraction	
	HgAc	2mM	20h	3	traitement impossible	
		4mM	20h	3,2	10,60%	29,90%
	Thimerosal	2mM	20h		pas de diffraction	
		2mM	17h	3,2	20,00%	32,40%
		0,5mM	60h	3,1	15,70%	39,60%
	Tetrakis	2mM	20h		pas de diffraction	
			5h		pas de diffraction	
	C ₂ H ₅ HgCl	2mM	20h		pas de diffraction	
5h			4,2	57,60%	41,30%	
Platine	KPtCl ₄	4mM	20h		pas de diffraction	
		2mM	20h		pas de diffraction	
	KPtCl ₆	4mM	20h	4	23,10%	32,30%
	NaPtCl ₆	2mM	20h	2,8	9,20%	20,70%
Or	Kau(CN) ₂	2mM	3h15	3,6	15,60%	23,50%
		4mM	20h	2,9	7,30%	20,80%
	KAuCl ₄	0,5mM	60h	3,6	28,20%	44,90%
Samarium	Sm(NO ₃)	2mM	20h	3,1	18,20%	42,50%
Gadolinium	Gd(NO ₃)	2mM	20h	3	9,80%	44,50%
Plomb	Pb(Ac)	2mM	20h	4,5	22,10%	73,20%
Tungstène	Na ₂ WO ₄	2mM	60h	3,5	8,50%	15,10%
Europium	Eu(NO ₃)	2mM	60h	3	traitement impossible	
Palladium	PdCl ₂	2mM	60h		peu de diffraction	
Barrium	BaCl ₂	2mM	60h		peu de diffraction	
Iode	KI	2mM	60h	4	15,90%	19,90%
Cobalt	CoCl ₂	2mM	60h		peu de diffraction	

Tableau 3.5 : Atomes lourds testés pour le phasage des cristaux d'eIF3i par la méthode MIR.

c Phasage par MAD

Principe

Nous avons considéré que la diffusion à l'origine de la diffraction consiste en une interaction élastique des photons X avec les électrons. Cette diffusion dite normale ne dépend alors pas de la longueur d'onde du rayonnement incident. En revanche, si l'interaction devient inélastique comme dans le cas d'une absorption, une partie de l'énergie incidente est absorbée et réémise sous forme de fluorescence. Le facteur de diffusion atomique, constant dans le cas précédent, change alors de valeur en fonction de la longueur d'onde et peut alors s'écrire sous la forme :

$${}^{\lambda}f = {}^0f + {}^{\lambda}f' + i{}^{\lambda}f''$$

L'introduction d'une composante imaginaire dans le facteur de diffusion atomique induit la rupture de la loi de Friedel, à savoir que $F_{hkl} \neq F_{\bar{h}\bar{k}\bar{l}}$. Cela fournit une information supplémentaire et des paires de réflexions : $F^+ (= F_{hkl})$ et $F^- (= F_{\bar{h}\bar{k}\bar{l}})$ sont ainsi déterminées.

La diffusion ainsi mesurée est dite anormale et ses variations sont étudiées aux alentours d'un seuil d'absorption où les variations des coefficients ${}^{\lambda}f'$ et ${}^{\lambda}f''$ sont les plus fortes. La technique de diffusion anormale à multiples longueurs d'onde (MAD) peut ainsi être mise à profit pour un dérivé d'atome lourd permettant ainsi d'éviter le recours à un deuxième dérivé pour lever l'ambiguïté sur les phases.

Plus généralement, ces enregistrements sont réalisés sur des cristaux d'une forme modifiée de la protéine correspondant à l'incorporation de sélénométhionine à la place des méthionines de la protéine. Dans ce cas, le signal anormal du sélénium est analysé pour déterminer les phases. L'enregistrement MAD débute par l'acquisition d'un spectre de fluorescence afin de déterminer la longueur d'onde correspondant au seuil d'absorption, qui va dépendre de l'environnement électronique des atomes fournissant le signal anormal. Cet enregistrement permettra ensuite d'évaluer les composantes de son facteur de diffusion atomique. En première approximation, le facteur ${}^{\lambda}f''$ se comporte comme la dérivée de ${}^{\lambda}f'$. La composante ${}^{\lambda}f'$ est minimale au point d'inflexion, ou seuil d'absorption, de la courbe (montante) ${}^{\lambda}f''$. La composante ${}^{\lambda}f''$ est quant à elle maximale au pic d'absorption.

La qualité du signal anomal peut être estimée par le rapport anomal entre les paires F^+ et F^- , dites paires de Bijvoët, à une même longueur d'onde et par le rapport dispersif calculé entre deux longueurs d'onde :

$$R_{ano} = \frac{\langle |F_{\lambda_1}^+| - |F_{\lambda_1}^-| \rangle}{\langle |F_P| \rangle} = \sqrt{\frac{N}{2}} \times \frac{2 \cdot \lambda_1 f''}{\langle |F_P| \rangle}$$

$$R_{disp} = \frac{\langle |F_{\lambda_1}| - |F_{\lambda_2}| \rangle}{\langle |F_P| \rangle} = \sqrt{\frac{N}{2}} \times \frac{|\lambda_1 f' - \lambda_2 f'|}{\langle |F_P| \rangle}$$

Ainsi le R_{ano} est maximal à la longueur d'onde du pic où $\lambda f''$ est maximal et R_{disp} est maximal entre la longueur d'onde de référence (loin du seuil d'absorption) et celle au seuil où $\lambda f'$ est minimal.

Le phasage peut alors être traité de façon similaire au cas du MIR en considérant chaque jeu de données correspondant à une longueur d'onde comme un jeu de données de dérivé. Afin de maximiser les différences entre les valeurs de $\lambda f'$ et $\lambda f''$, trois longueurs d'ondes sont en général utilisées : la longueur d'onde au seuil, celle du pic et une longueur d'onde de référence, éloignée du seuil. Les programmes de phasage solve et sharp peuvent être utilisés pour le phasage par MAD. Le signal anomal est généralement faible et requiert d'enregistrer les paires de Bijvoët sur des clichés assez rapprochés dans le temps afin d'éviter notamment les problèmes de mise à l'échelle dues à la décroissance du faisceau de rayons X. En revanche, le défaut d'isomorphisme est généralement inexistant dans ce système, ce qui permet d'obtenir des phases de départ de très bonne qualité, au regard de celles obtenues en MIR.

Obtention de protéine sélénométhionylée et phasage

La bioincorporation de sélénométhionine est réalisée en transformant une souche d'*E. coli* auxotrophe pour la méthionine : B834(DE3). Suivant une préculture de ces bactéries transformées en milieu riche, les cellules sont lavées dans un milieu minimum. Ce milieu minimum ne contient pas de méthionine mais de la sélénométhionine. Alternativement, l'utilisation de milieux autoinductibles permet d'utiliser une souche sans auxotrophie ([Studier, 2005](#)). Ces milieux contiennent un ratio glucose/lactose assurant la consommation du glucose durant la phase exponentielle et donc la répression de la traduction du gène porté par le plasmide. En fin de phase exponentielle, lorsque le glucose est consommé, le recours au lactose induit la production de la protéine. Notamment, le milieu adapté à la production de protéines sélénométhionylées contient un ratio de méthionine/sélénométhionine limitant la toxicité liée à la présence du deuxième composé tout en assurant une bioincorporation de sélénométhionine de 80-90%. Clairement, la seconde technique permet d'augmenter la surproduction puisqu'à partir d'un litre de culture de B834, environ 5mg de protéine sont élués de la résine TALON alors que 250mL de milieu autoinductible permettent d'en obtenir environ 25mg. Néanmoins, dans les deux cas, la coupure de l'étiquette polyhistidine induit la précipitation de la protéine, que cette étape soit réalisée lors d'une dialyse longue (une nuit), courte (30min. à 1 heure) ou en « batch ». Systématiquement, s'il reste de la protéine, celle-ci reste accrochée sur la résine TALON et il n'a jamais pu être obtenu de protéine sans étiquette. De façon comparable à la protéine native, la cristallisation n'a pas pu être obtenue lorsque la protéine comporte l'étiquette.

Compte tenu de l'impossibilité d'enlever l'étiquette N-terminale d'eIF3i, nous avons construit un nouveau plasmide permettant l'étiquetage C-terminal de la protéine ([voir Matériels et Méthodes](#)). Une très bonne surproduction de la protéine native a été obtenue sous cette forme (environ 200mg élués de la résine IMAC par litre de culture), encore meilleure que celle obtenue avec la forme d'eIF3i fusionnée en N-terminal (50mg par litre de culture). La coupure à la thrombine est efficace et le même protocole de purification a pu être suivi. De la même façon, la forme sélénométhionylée de cette protéine a pu être obtenue et purifiée de manière identique. Cependant, autant pour la protéine native que pour la protéine sélénométhionylée, aucune cristallisation n'a pu être obtenue, que ce soit dans les différentes matrices de conditions de cristallisation ou celle identifiée pour la protéine obtenue avec la fusion N-terminale d'une polyhistidine.

Enfin, la production de la protéine sélénométhionylée sans étiquette à partir d'un vecteur pET3a n'a pas permis de purifier la protéine.

Les différences observées entre les fusions N et C-terminales pourraient être liées à la structure de la protéine. Il faut noter que la coupure de l'étiquette polyhistidine grâce à la thrombine laisse 2 résidus supplémentaires par rapport à la séquence native d'eIF3i. Or, le dernier motif lame des repliements en hélices β est composé de brins issus de la partie N-terminal sur lesquels se replie un brin de l'extrémité C-terminale accordant aux résidus terminaux une grande importance pour le maintien de la structure (**Figures 3.19 et 3.20**).

Dans ce contexte, en alignant les séquences d'eIF3i issues de différents organismes, il peut être remarqué que les séquences d'eIF3i sont très conservées à leur extrémité N-terminale et que le facteur de levure comporte une extension C-terminale de 12 résidus par rapport aux formes les plus courtes, et notamment au facteur humain.

Il a donc été tenté de produire trois versions tronquées de la protéine :

La première consiste à supprimer des résidus C-terminaux afin que la protéine sans étiquette fasse la même longueur que tout un groupe de facteur d'eIF3i comportant 5 résidus additionnels par rapport au facteur humain (**Figures 3.22a, b, facteurs encadrés en vert**). Deux formes ont été clonées, de manière à exprimer la protéine fusionnée à une étiquette en N-terminal d'une part et en C-terminal d'autre part.

Une seconde version permet de produire une protéine comportant une étiquette en C-terminal qui, une fois hydrolysée par la thrombine, contient le même nombre de résidus que la forme d'eIF3i de levure pour laquelle nous avons obtenu la cristallisation.

Une dernière version concerne la fusion N-terminale et permet de produire une protéine ayant le même nombre de résidus C-terminaux que le facteur humain (**Figures 3.22a, b, facteur encadré en orange**).

Les vecteurs exprimant chacune de ces formes ont été obtenus par mutagenèse de délétion pour la fusion C-terminale et par mutagenèse ponctuelle en substituant le codon C-terminal par un codon STOP pour la fusion N-terminale. Ces protéines sous forme natives ou sélénométhionylées ont pu être purifiées jusqu'à homogénéité. Il n'a cependant pu être obtenu de cristallisation que ce soit avec la condition précédemment identifiée ou à l'aide des matrices de solutions de cristallisation.

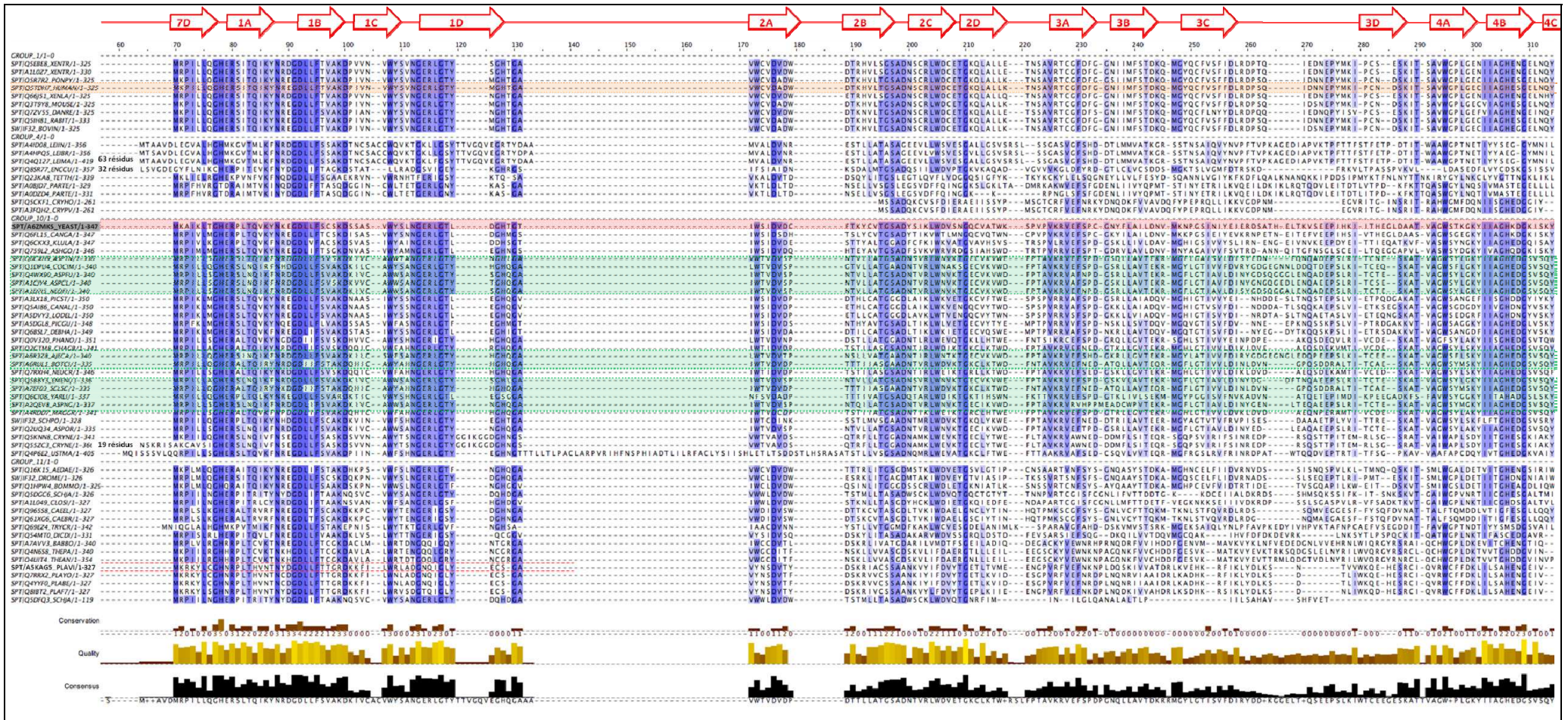


Figure 3.22a : Alignement de séquences d'eIF3i réalisé avec PipeAlign (Plewniak et al., 2003). La conservation des résidus est représentée par une échelle de bleu, le bleu foncé représentant une forte conservation. Les protéines homologues ont été sélectionnées et groupées en fonction des homologies par le programme Ballast (Plewniak et al., 2000). La séquence d'eIF3i de levure est encadrée en rouge et la séquence du facteur humain apparaît en orange. Les facteurs du groupe d'eIF3i de levure possédant une extension C-terminale de 5 résidus, par rapport au facteur humain, sont indiqués en vert.

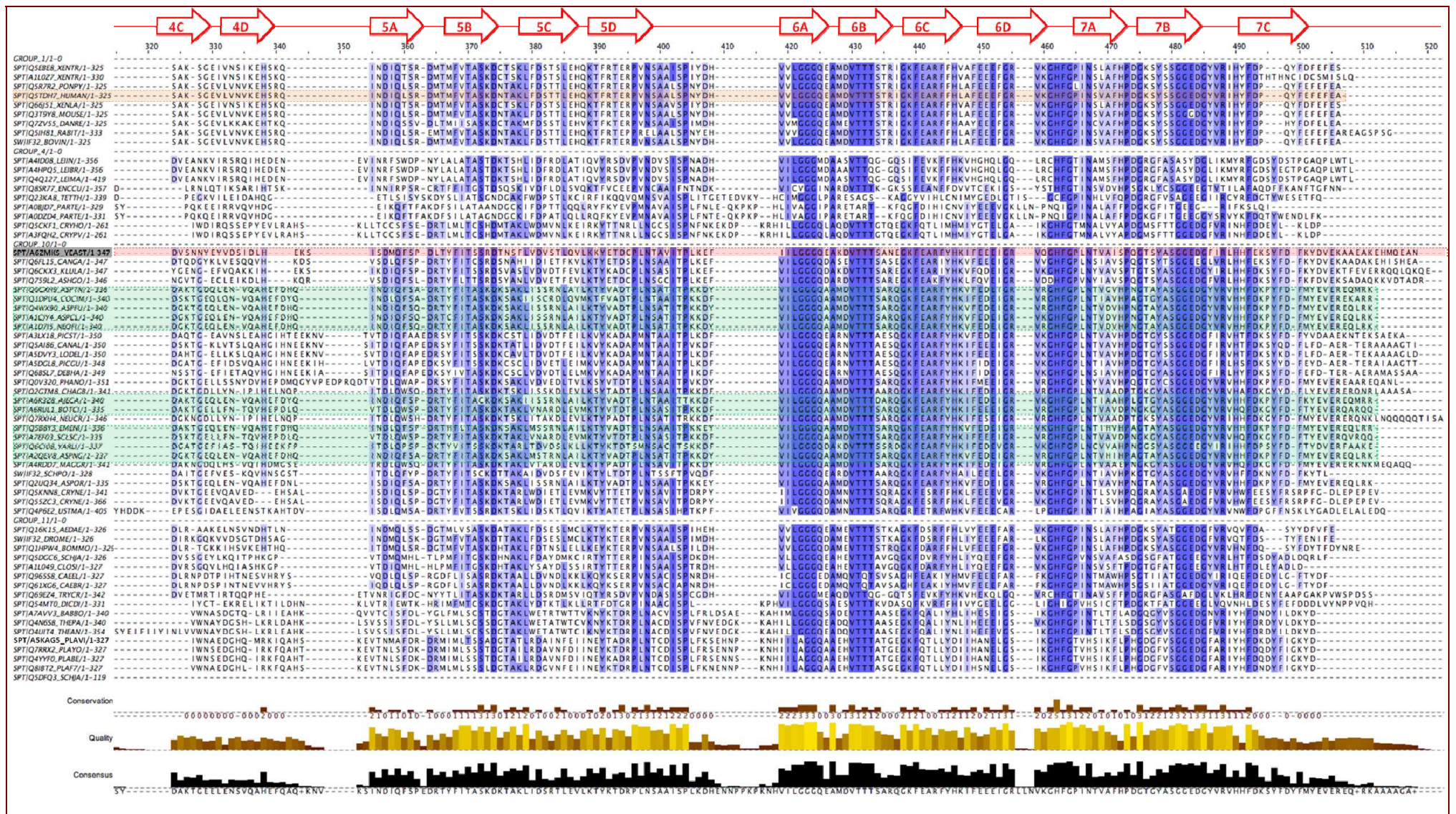


Figure 3.22b : Alignement de séquences d'eIF3i réalisé avec PipeAlign (Plewniak et al., 2003). Voir Figure 3.22a

DISCUSSION

La résolution de la structure du facteur eIF3i de *S. cerevisiae* s'avère représenter un défi à deux niveaux. Premièrement, les cristaux sont difficiles à obtenir et cela est reflété par l'impossibilité au cours de ce travail de produire des cristaux de la forme sélénométhionylée. La cocrystallisation de la protéine en présence d'atomes lourds représente une dernière piste que nous souhaitons explorer.

La seconde limitation provient des cristaux en eux-mêmes et de leurs caractéristiques. Le jumelage combiné à la NCS de translation pure présente un problème cristallographique difficile. La résolution de structure à partir de cristaux jumelés n'est cependant pas théoriquement impossible. Il apparaît néanmoins que la détermination des phases soit dans ce cas plus aisée en utilisant la méthode de remplacement moléculaire. Les techniques expérimentales de phasages permettent d'ajouter aux phases issues du remplacement moléculaire des informations afin de les préciser. Les techniques MIR et le MAD sont sujettes à produire également des phases initiales mais l'étape de « déjumelage » souvent requise représente une étape critique qui peut aboutir à la perte d'information. Une implémentation annoncée du programme phaser permettra de traiter la NCS de translation lors de la recherche de la translation. Ceci constitue une grande attente puisqu'il apparaît que la recherche simultanée des deux molécules permette de contourner le problème de superposition des intensités liées au jumelage.

Finalement, pour pallier les difficultés liées au facteur eIF3i de levure nous envisageons de tester la production de facteurs eIF3i issus d'autres organismes et notamment le facteur humain et celui d'*Encephalitozoon cuniculi*.

Plus généralement, le développement d'une stratégie de purification basée sur un système de coexpression nous a permis de considérablement améliorer les rendements des purifications tout en limitant les opérations menant à la reconstitution de sous-complexes. Très rapidement, la coexpression a fait ses preuves sans nécessiter d'adaptation de l'organisation des différents ORF au sein des plasmides. En outre, une série d'échange avec différents laboratoires (IGBMC (Strasbourg), IBMC (Strasbourg)) autour du thème des « Stratégies de purification de complexes macromoléculaires » a permis de confirmer les évolutions que nous avons apportées à notre système de purification. Notamment, l'expression altérée d'eIF3a que nous observons lors de la production d'un complexe intact ne semble pas être un cas isolé lors de la mise en place d'une stratégie de coexpression. Dans cette situation, il apparaît intéressant de combiner l'approche en coexpression avec une reconstitution *in vitro* du complexe. Après un travail d'affinement de la stratégie, nous devrions être en mesure de produire un complexe de 323kDa formé de toutes les sous-unités du facteur de démarrage eIF3 sur lequel une étude structurale pourra être entreprise une fois une stœchiométrie acceptable obtenue.

En conclusion, à mon arrivée au laboratoire, nous ne disposions que des plasmides susceptibles de produire les sous-unités isolées. La littérature avait permis de créer une vue « artistique » du facteur eIF3 de levure ([Valasek et al., 2002](#)). Une structure par cryomicroscopie électronique a permis de visualiser la forme du facteur humain mais la purification du complexe directement à partir d'un extrait brut de cellules ne permet pas d'obtenir plus d'informations sur l'organisation du complexe eIF3 ([Siridechadilok et al., 2005](#)).

Mes travaux ont abouti à la construction d'une grande bibliothèque de vecteurs utilisés pour purifier l'ensemble des sous-complexes qui devraient permettre de décrire eIF3 dans son entier, par microscopie électronique et/ou par cristallographie aux rayons X.

Chapitre IV Matériels et Méthodes

I Techniques générales de biologie moléculaire, réactifs et tampons de purification

1 Souches de bactéries

Les différentes souches de bactéries ont été transformées par électroporation effectuée dans un appareil Gene Pulser (Biorad, Hercules, Californie) selon les recommandations du constructeur. Les bactéries électrocompétentes ont été préparées selon le protocole de Dower (1988). Les génotype des différentes souches utilisées dans cette étude sont répertoriées dans le **tableau 4.1**.

Souche	Génotype	Référence
JM101Tr	$\Delta(lac-pro) supE thi recA56 srl-300::Tn10$ (F' <i>traD36 proAB lacI^q lacZ</i> Δ M15)	Hirel (1988)
XL1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1</i> <i>lac [F'proAB lacI^q lacZ</i> Δ M15 Tn10 (TetR)]	Stratagene
BL21(DE3)	F ⁻ <i>ompT hsdS_B (r_B⁻ m_B⁻) gal dcm λ(DE3)</i>	Novagen
Rosetta(DE3)pLacI	F ⁻ <i>ompT hsdS_B (r_B⁻ m_B⁻) gal dcm λ(DE3)</i> pLacIRARE (CmR)	Novagen
B834(DE3)	F ⁻ <i>ompT hsdS_B (rb- mB-) gal dcm λ(DE3) met</i>	Novagen

Tableau 4.1 : Génotype des souches utilisées dans cette étude

Les compositions des différents milieux de culture utilisés sont fournis ci-après pour 1L, les différents composants sont commandés chez Difco :

Milieux liquides :

LB : 10g de tryptone, 5g d'extrait de levure, 5g de chlorure de sodium (NaCl)

2xTY : 16g tryptone, 10g d'extrait de levure, 5g de NaCl

M9 (x5) : 37,5g Na₂HPO₄ ; 15g KH₂PO₄ ; 2,5g NaCl ; 6,25g NH₄Cl

Milieux solides :

LBAgar : Milieu LB + 0,2% w/v d'agar

HTOP : 10g tryptone, 8g NaCl, 8g agar

Les quantités d'antibiotiques utilisés pour sélectionner les résistances sont indiquées dans le **tableau 4.2**.

Antibiotique	Abréviation	Solution (mg/mL)	Utilisation (mL/L)	Concentration finale (µg/mL)
Chloramphénicol	Cm	34	1	34
Ampicilline	Amp	25	4	50
Kanamycine	Kan	25	1	25
Spectinomycine	Spe	20	1	20
Streptomycine	Sm	20	1	20

Tableau 4.2 : Concentrations d'antibiotique utilisées pour les sélections

2 Plasmides, préparations plasmidiques et caractérisation

Les plasmides pET15b1pa et pET3alpa ont été obtenus par l'insertion dans les plasmides correspondants respectifs d'une séquence entre les sites *NdeI* et *BamHI* correspondant les sites de restrictions des enzymes *SmaI*, *XhoI*, *KpnI*, *ScaI*, *NotI* et *SaI* (TATGCCCGGGCTCGAGGGTACCCCGCGGGCGGCCGCGTCGA CC). Le plasmide pET15bXSB a été obtenu par ajout d'un site *SmaI* entre les sites *XhoI* et *BamHI* du plasmide pET15b.

Les minipréparations d'ADN plasmidiques ont été obtenues selon le protocole décrit par Birnboim et Doly ([Birnboim and Doly, 1979](#)) ou par l'utilisation d'un protocole automatisé opéré par le robot Genesis (Tecan) en utilisant la méthode Nucleobond Spin (Macherey Nagel). Les systèmes QIAfiler (QIAGEN) ou Nucleobond (Macherey Nagel) ont permis d'obtenir des préparations d'ADN plasmidiques à grande échelle.

La caractérisation par électrophorèse d'ADN ont été réalisées par migration de gels à 1% p/v d'agarose LE (Roche, Manheim, Allemagne) en présence de bromure d'éthidium à 0,3mg/mL (AMRESCO). Les migrations sont réalisées dans un tampon TBE {Tris 50mM, Borate 60mM, EDTA 1mM, pH 8,3}.

Les concentrations des préparations d'ADN plasmidique sont évaluées par l'acquisition de spectres d'absorption dans la gamme UV-visible sur un appareil UC2101PC (Shimadzu) ou ND-1000p (NanoDrop) en considérant qu'une unité d'absorbance à 260nm correspondant à une valeur approximative de 50 µg/mL d'ADN double brin.

3 Clonage

Les manipulations des séquences d'ADN plasmidique ont été effectuées à l'aide du logiciel DNAid ([Dardel and Bensoussan, 1988](#)). Les oligonucléotides utilisés ont été commandés chez MWG Biotech AG (Ebersberg, Allemagne) et les désoxynucléotides proviennent d'Amersham (Uppsala, Suède). Les enzymes de restrictions ont été commandées chez Roche (Manheim, Allemagne) excepté pour les enzymes DpnI et HindIII provenant de New England Biolabs (Beverly, Massachusetts).

L'amplification des fragments d'ADN par PCR ont été réalisés dans un volume réactionnel de 25µL contenant 5 ou 10ng de matrice plasmidique ou 10ng de matrice génomique, 50pmoles de chacune des amorces nucléotidiques et un concentration finale de 200µM en chacun des désoxyribonucléotides (dATP, dCTP, dGTP, dTTP) ainsi que 2 unités d'ADN polymérase thermosensible Taq DNA Polymerase (Roche) et 0,4 unités de la polymérase Cloned Pfu (Stratagen) dans le tampon fourni avec la Taq DNA Polymerase. L'utilisation de la polymérase de Stratagene permet une relecture du fragment amplifié et donc une meilleure fidélité de cette biosynthèse. Les réactions de PCR sont réalisées en 30 cycles précédés d'une phase préalable de dénaturation à 95°C pendant 2min. Les cycles comportent alors 15s de dénaturation à 95°C, 30s d'hybridation des oligonucléotides sur la matrice d'ADN (50-55°C en fonction des oligonucléotides) et une minute d'élongation à 72°C. Le programme d'achève par une phase d'élongation de 5min. à 72°C. La quantité de fragment synthétisé est alors contrôlée par dépôt de 5µL du mélange réactionnel sur gel d'agarose. Le mélange réactionnel est alors injecté sur un tamis moléculaire adapté à la purification de l'ADN (TSK Gel DNA, TOSOHAAS, Japon) ([Schmitter et al., 1986](#)). Les fractions de 320µL correspondant aux fragments d'ADN sont concentrés par précipitation en présence d'éthanol (EtOH) à 66%, d'acétate de sodium (NaAc) pH 4,6 à 0,3M et 10µg de Dextran DT40 par fraction. Ils sont ensuite hydrolysés par une quantité adaptée des enzymes de restriction avant d'être précipités par un mélange EtOH/NaAc. Les vecteurs de destination sont de la même façon hydrolysés par les mêmes enzymes de restriction avant d'être précipité par le même mélange EtOH/NaAc pH4,6 et dissouts dans l'eau. Les quantités d'ADN sont ensuite évaluées par électrophorèse afin de réaliser la ligature en présence d'un excès molaire de 7 à 10 de fragment à insérer par rapport au vecteur.

Les ligatures ont été réalisées avec l'ADN ligase du phage T4 (GIBCO BRL) ou l'ADN ligase FastLink (TEBU Bio, Epicentre Technologies, Madison, Wisconsin). Lorsque l'ADN ligase du phage T4 est utilisée, un mélange réactionnel de 20-25µL contenant 10ng de vecteur, la quantité appropriée de fragment digéré à insérer, 200µM final d'ATP, 4mM final de DTT, 2,5 unité de ligase en présence de 50mM Tris-HCl pH 8,0, 10mM MgCl₂. Le mélange est alors laissé une heure sur paillasse puis la ligase est inactivée en chauffant le mélange réactionnel pendant 10 minutes à 70°C. Les réactions de ligatures sont alors hydrolysées par une enzyme de restriction dont un site est présent sur le vecteur de départ et pas sur le plasmide combiné attendu. Si nécessaire, une précipitation par un mélange EtOH/NaAc pH 4,6 est réalisée pour placer l'ADN dans un tampon compatible avec l'enzyme de restriction. Dans certains cas, le vecteur de départ digéré est purifié sur la colonne TSK Gel DNA afin d'être séparé du fragment généré par la digestion et limitant l'efficacité de la ligature.

1µL du mélange réactionnel est utilisé pour transformer les bactéries de la souche JM101Tr. Suivant l'électroporation, les bactéries sont mélangées à 3mL de milieu HTOP préchauffé à 40°C avant d'être étalées sur des boîtes de Pétri contenant un milieu LB Agar complémenté par un antibiotique adapté à la résistance du plasmide. Après une nuit de culture à 37°C, 8 colonies par clonages sont repiquées sur des boîtes LB Agar complémentées en antibiotique et cette procédure est répétée le lendemain afin d'assurer la clonalité des colonies. Les minipréparations d'ADN plasmidiques sont ensuite réalisées. L'insertion est vérifiée par analyse de restriction. Deux clones identifiés comme positifs servent à ensemercer 100mL de milieu LB complémenté en l'antibiotique adapté afin d'extraire l'ADN plasmidique en plus grande quantité. La séquence de l'insert est systématiquement déterminée avant que le plasmide ne serve à l'expression de la protéine codée par le fragment inséré.

4 Mutagenèse dirigée

Les mutants ponctuels ainsi que les mutations par insertion/délétion ont été réalisées par la méthode QuickChange™ Site-Directed Mutagenesis (Stratagene). Les mutations sont obtenues à l'aide d'oligonucléotides dont les extrémités 5' et 3' s'hybrident sur le vecteur alors que la partie interne comporte la séquence que l'on veut insérer à cette position. A l'inverse, les mutations de délétion consistent à créer une boucle au niveau du vecteur par l'hybridation d'un oligonucléotide complémentaire aux extrémités de la partie du gène à supprimer. L'ADN polymérase se basant sur l'amorce nucléotidique pour sa synthèse ne recopiera pas la partie du vecteur formant la boucle.

Le protocole alors utilisé consiste à préparer 25µL d'un mélange de 50 ou 100ng d'ADN cible, 15 pmoles de chacune des amorces, 200µM de chacun des quatre désoxynucléotides, 2,5 unités de la polymérase PfuTurbo dans un tampon fourni. La mutagenèse est alors réalisée par respectivement 12 ou 18 des cycles suivants pour des mutations ponctuelles ou des mutations par insertion/délétion. Les cycles sont précédés d'une étape préliminaire de dénaturation pendant 2min. à 95°C. Chaque cycle comporte alors 30s de dénaturation à 95°C, 30s d'hybridation à 50°C et 10min. d'élongation à 68°C. La durée de la phase d'élongation et la température d'hybridation ont été optimisés en fonction de la taille de la matrice à recopier et des caractéristiques des oligonucléotides. La synthèse se termine par une phase additionnelle d'élongation de 10 minutes à 68°C. La qualité de l'ADN obtenu est alors contrôlée par dépôt sur gel d'électrophorèse de 10µL du mélange réactionnel.

Il s'ensuit une hydrolyse d'une heure par 10 unités de l'enzyme *DpnI* qui a la capacité de couper spécifiquement les brins méthylés de l'ADN. Ainsi, puisque seul le plasmide de départ comporte ces modifications qui ont lieu *in vivo*, il sera hydrolysé alors que le plasmide synthétisé *in vitro* ne sera pas modifié. Les hétéroduplexes sauvage/mutant sont également hydrolysés par *DpnI*. 1µL du mélange réactionnel est alors utilisé pour transformer des bactéries XL1 Blue qui sont déficientes en l'enzyme de restriction *hsdR*, capable d'identifier de l'ADN non-méthylé et donc induirait la destruction du plasmide synthétisé *in vitro* inséré dans la bactérie. La transformation est alors mélangée à 3mL de HTOP préchauffé avant d'être réparti sur une boîte contenant un milieu LB agar permettant de sélectionner la résistance du plasmide muté. Après deux étapes d'isolement de colonies sur boîtes de Pétri contenant l'antibiotique adapté, des minipréparations d'ADN plasmidiques sont entreprises pour 8 colonies. La séquence des ADN obtenus est réalisée afin d'identifier les clones mutés qui serontensemencés dans 100mL de milieu LB sélectif pour préparer une plus grande quantité de plasmide. Les deux clones retenus seront à nouveau séquencés. Dans le cas des mutagenèses d'insertion/délétion, des restrictions des minipréparations plasmidiques sont entreprises lorsque la taille du fragment inséré/supprimé est suffisante. Le séquençage des secondes préparations d'ADN de deux clones positifs est ensuite effectué.

5 Séquençage de l'ADN

Les séquences des différents plasmides ont été vérifiées en utilisant un séquenceur automatique Long ReadIR 4200 (Li-Cor, Lincoln, Nebraska). Les gels de migration ainsi que le tampon de migration ont été préparés en accord avec les recommandations du constructeur pour réaliser des gels de 41cm. Les fragments d'ADN obtenus par la méthode de terminaison de chaîne décrite par Sanger ([Sanger et al., 1982](#)) sont détectés par excitations de groupements fluorophores dans l'infrarouge (IRDye 700 et IRDye 800), greffés à l'extrémité 5' d'amorces nucléotidiques de 21 bases commandées chez MWG Biotech AG (Ebersberg, Allemagne). Les réactions de séquences par cyclage thermique ont été réalisées à l'aide de la méthode de séquençage d'Epicycle Technologies (TEBU Bio, Madison, Wisconsin) dans les appareils GeneAmp 2400 (Perkin Elmer, Wellesley, Maryland). Les séquences sont déterminées par le logiciel du séquenceur et les électrophorégrammes utilisés pour lever les ambiguïtés. Les alignements de séquences d'ADN ont été réalisés à l'aide du programme DNASTrider ([Marck, 1988](#)).

6 Caractérisation des protéines par SDS-PAGE

La séparation des protéines en fonction de leurs tailles a été réalisée par électrophorèse en conditions dénaturantes (SDS-PAGE) à l'aide du système MiniProtean II de Biorad (Hercule, Californie). Les gels et tampon de migration ont été préparés selon Laemmli et al. ([Laemmli et al., 1970](#)). Suivant la migration, les bandes de protéine sont révélées par coloration au bleu de Coomassie (Coomassie Brilliant Blue de SIGMA) ou par coloration au nitrate d'argent (Silver Stain Plus de Biorad).

7 Caractérisation des protéines par électrophorèse bidimensionnelle

Une caractérisation plus poussée des complexes peut-être fournie par une migration électrophorétique bidimensionnelle en séparant dans un premier temps les protéines en fonction de leurs points isoélectriques (pI) puis selon leur taille. La première migration a été réalisée en utilisant des fines bandes de gels (bandes tampon) commercialisés par Biorad couvrant une gamme de pI entre 4 et 7 dans un appareil Protein IEF Cell (Biorad). Les bandes de gels issues de la première migration sont ensuite déposées sur un gel classique pour la SDS-PAGE et la seconde migration réalisée dans les unités MiniProtean II pour des bandes de 7cm et Protean II pour les bandes de 17cm. La révélation est ensuite réalisée par coloration au bleu de Coomassie.

8 Caractérisation des protéines par électrophorèse en condition native

La détermination de la composition des tampons de séparation, de dépôt et de migration sont déterminées en fonction des caractéristiques physico-chimiques des protéines et notamment leurs points isoélectriques. Dans le cas des facteurs archéens, leurs points isoélectriques basiques imposent une électrophorèse en conditions natives acides alors que les conditions basiques sont destinées aux protéines acides de la levure. L'analyse des complexes protéiques par électrophorèse en conditions natives sont réalisés avec le même système d'électrophorèse MiniProtean II. La composition des tampons de migration et de ceux utilisés pour préparer les gels est reportée dans le **tableau 4.3**. Les gels sont préparés de façon identique à ceux des conditions dénaturantes en omettant le

SDS. Les migrations de gel dans ce cas sont réalisées en chambre froide à 4°C, à intensité contrôlée (50mA) pour assurer l'intégrité des complexes au cours de l'électrophorèse. La procédure de révélation des bandes au bleu de Coomassie est similaire à celle décrite pour les gels SDS-PAGE.

Conditions acides				Conditions basiques		
Tampon	Séparation (x4)	Dépôt (x4)	Migration (x5)	Séparation (x4)	Dépôt (x4)	Migration (x5)
Base	48mL KOH 1M	48mL KOH 1M	31,2g alanine	27,2g Tris	6g Tris	30,29g Tris
Acide	17,2mL ac. acétique	2,9mL ac. acétique	8mL ac. acétique			7,73g ac. Borique
Eau	34,8mL	49,1mL	Q.S.P. 1L	Q.S.P. 150mL	Q.S.P. 100mL	Q.S.P. 1L
pH	ajusté à 4,3 par KOH	ajusté à 6,8 par KOH	4,4	ajusté à 8,8 par HCl	ajusté à 6,8 par HCl	8,7

Tableau 4.3 : Composition des tampons utilisés pour l'électrophorèse

9 Cocktail antiprotéases

Afin de limiter la protéolyse des protéines surproduites, un mélange d'antiprotéases est utilisé lors du cassage des parois cellulaires. Le **tableau 4.4** indique la composition ainsi que la provenance des différents composants de ce cocktail concentré 50 fois par rapport à sa concentration d'utilisation et noté CAPx50.

Composé	Concentration finale (µg/mL)	Pour 100mL (mg)
Aprotinine	25	2,5
Antipaïne	25	2,5
Chymostatine	25	2,5
Leupeptine	25	2,5
Pepstatine A	25	2,5 + 1mL EtOH + 222µL ac. acétique
Benzamidine	5000	500
Phenantroline	50	5 + 500µL EtOH
Ovonucoïde	30	30

Tableau 4.4 : Composition du cocktail antiprotéase CAPx50

10 Purification à grande échelle des protéines

Ces purifications ont été réalisées sur les systèmes FPLC Liquid Chromatography Controller LCC-500 (Amersham Pharmacia Biotech, Upssala, Suède) ou bien par le système chromatographique de Biorad assisté par ordinateur BioLogic DuoFlow. Les colonnes de chromatographie Q-Hiload, S-Hiload, Mono Q, Mono S, Superdex 75, Superdex 200, Superdex 75HR, Superdex 200HR, HIC-C3 proviennent du fabricant Amersham Pharmacia Biotech. La résine IMAC TALON permettant l'accrochage spécifique des étiquettes histidine est commandée chez Clonetech. La protéase thrombine utilisée pour couper l'étiquette His₆ est commandée chez Sigma.

Les protéines purifiées ont été concentrées par centrifugation sur membrane en utilisant des unités de centrifugation Vivaspin 20mL, 10000MWCO ; Vivaspin 500µL, 5000MWCO (Vivascience AG, Hannover, Allemagne) ou les unités AmiconUltra 4mL, 10000 ou 30000MWCO (Millipore).

Les coefficients d'extinction molaires des protéines sont calculés à l'aide du logiciel DNAid. Les concentrations de protéine sont ainsi estimées par l'enregistrement de spectre d'absorption dans la gamme UV-visible sur un appareil UV2101PC ou ND-1000.

Le cassage des parois cellulaires est réalisé par sonication des bactéries resuspendues dans 50mL de tampon adapté (tampon de cassage).

La composition des tampons utilisés au cours des différentes purifications sont indiquées ci-dessous :

Tampons de cassage :

C₁ {10mM MOPS pH 6,7 ; 10mM 2-mercapto-éthanol ; 200mM NaCl ; 0,1mM PMSF}

C₂ {10mM MOPS pH 6,7 ; 3mM 2-mercapto-éthanol ; 500mM NaCl ; 0,1mM PMSF}

Tampons de séparation :

A₁ {10mM MOPS pH 6,7 ; 10mM 2-mercapto-éthanol ; 0,1mM PMSF}

B₁ {10mM MOPS pH 6,7 ; 10mM 2-mercapto-éthanol ; 1M NaCl ; 0,1mM PMSF}

A₂ {10mM MOPS pH 6,7 ; 3mM 2-mercapto-éthanol ; 0,1mM PMSF}

B₂ {10mM MOPS pH 6,7 ; 3mM 2-mercapto-éthanol ; 1M NaCl ; 0,1mM PMSF}

SA {10mM MOPS pH 6,7 ; 3mM 2-mercapto-éthanol ; 1,7M SA ; 0,1mM PMSF}

Tampon de digestion par la thrombine :

T {10mM MOPS pH 6,7 ; 3mM 2-mercapto-éthanol ; 500mM NaCl ; 10mM CaCl₂}

11 Essais de surproduction des protéines comportant une étiquette His₆

Préalablement à toute purification à grande échelle, des essais de purification des formes de protéines fusionnées à une étiquette His₆ correspondant aux gènes nouvellement clonés sont réalisés. Pour cela, la souche Rosetta(DE3)-pLacIRARE™ est transformée par le plasmide correspondant. 50µL des bactéries reprises dans 1mL de 2xTY sont utilisées pour ensemercer 50mL de 2xTY complété par les antibiotiques adaptés à la sélection de la souche et du plasmide. La culture est ensuite induite par ajout d'IPTG à 1mM final à différentes phases de la croissance cellulaire ; en général, au plateau et en phase exponentielle. Suivant 3-4 heures d'induction, les bactéries sont recoltées par centrifugation.

Les bactéries sont suspendues dans 1,5 mL de tampon C₂ complété par 1/50ème v/v de CAPx50. Les parois cellulaires sont ensuite rompues par sonication pendant 120s à une puissance la plus élevée possible. On procède par impulsions, séparées par des pauses identiques à celles des impulsions, pour permettre le refroidissement de l'échantillon. Le mélange est centrifugé pour séparer les débris cellulaires de la partie soluble (30min., 15300 rpm, 4°C). Le mélange est alors laissé en contact à 4°C avec 20-30µL de résine TALON préalablement équilibrée dans l'eau puis le tampon C₂. Suivant les 40 minutes d'adsorption sur la résine, celle-ci est disposée dans les micro-colonnes d'1,5mL MicroBioSpin (Biorad). L'éluat est alors recueilli par gravitation. Suivant 3 rinçages par le tampon C₂, les protéines accrochées à la résine sont éluées par 10-20µL de tampon C₂ complété par une concentration de 250mM finale en imidazole (Sigma). Des échantillons correspondant à la sonication, au culot de sonication resuspendu, à l'éluat primaire de la colonne (Flow Through : FT), au dernier lavage ainsi que l'éluat sont finalement analysés par SDS-PAGE.

12 Digestion de bandes excisées de gels pour analyse par SM

Le protocole suivi pour l'excision des bandes et l'hydrolyse par la trypsine a été développé et est régulièrement implémenté par l'équipe de la Plateforme Spectrométrie de masse biologique et Neuroprotéome (ESPCI, Paris). Le protocole est disponible en ligne (<http://www.bio.espci.fr/spectro/protocoles.php>).

13 Autres produits et réactifs

Les autres produits chimiques utilisés au cours de cette étude ont été commandés chez Merck, Fluka ou Sigma-Aldrich.

II Etude du facteur e/aIF5B

1 Clonage et Purification de ss-aIF5B

Le gène codant pour la protéine aIF5B a été amplifié par PCR à partir de l'ADN génomique de *Sulfolobus solfataricus* puis cloné entre les sites *NdeI* et *SacII* du plasmide pET3alpa permettant d'obtenir pET3a5Bss.

Pour produire la protéine, des bactéries Rosetta(DE3)-pLacIRARE sont transformées par ce plasmide puis mises en cultures d'1L du milieu 2xTY complété par les antibiotiques Amp et Cm à 37°C. Après une nuit de culture, l'induction des bactéries à une DO_{260nm} d'environ 1,6 est réalisé par ajout d'IPTG à une concentration finale de 1mM et la culture est prolongée de 5 heures à température ambiante. Les cellules sont lysées par sonication dans 40mL de tampon C₁. Après centrifugation (30min. à 15300rpm), le surnageant est chauffé à 70°C pendant 10min. Les protéines précipitées sont éliminées par centrifugation (10min., 15300rpm). Le surnageant est soumis à une précipitation au sulfate de streptomycine à 1% w/v final et le mélange clarifié par centrifugation (30min., 15300rpm). Le surnageant est enfin précipité au sulfate d'ammonium jusqu'à 80% de saturation. Le culot est resuspendu dans 4mL de tampon A₁10%B₁ puis dialysé 1 heure contre ce même tampon. L'échantillon est ensuite appliqué sur une colonne Q-Hiload équilibrée dans ce tampon. La protéine n'accroche pas sur cette résine. Les fractions correspondantes sont diluées 3 fois dans le tampon A₁ avant d'être chargées sur une colonne S-Hiload. L'application d'un gradient de 10 volume de colonne de A₁3,3%B₁ à A₁50%B₁ permet l'élution de la protéine. Les fractions associées sont concentrées à 3mL afin d'être appliquées sur une colonne Superdex75 équilibrée dans le tampon A₁50%B₁. La protéine est éluée sous un pic unique et sous forme homogène comme l'indique une analyse SDS-PAGE. La protéine est finalement concentrée à 7mg/mL. La concentration est évaluée par mesure de densité optique en considérant un coefficient d'extinction molaire calculé à 0,437cm²/mg. 15mg de protéines sont ainsi obtenus à partir d'un litre de culture.

2 Clonage et Purification des domaines de ss-aIF5B

Le gène correspondant au domaine IV du facteur a été obtenu par l'ajout d'un site *NdeI* au niveau du codon TTT de l'acide aminé F462 introduisant ainsi un codon ATG. Le fragment amplifié par PCR a ensuite été cloné entre les sites *NdeI* et *XhoI* du plasmide pET3alpa. Le vecteur exprimant les domaines I, II et III de ss-aIF5B résulte de la mutation d'un codon GAA du résidu 455 par un codon STOP (TGA).

Les protéines sont exprimées et purifiées de la même façon que la forme native excepté que le domaine IV est dialysé une heure avant l'application sur la colonne S-Hiload. Les concentrations en protéines sont évaluées à l'aide des coefficients d'extinction molaire calculés de 0,378 et 0,463cm²/mg pour les domaines IV et I, II et III respectivement. A partir d'un litre de culture, 2mg du domaine IV et 10mg du domaine I, II et III ont ainsi pu être purifiés jusqu'à homogénéité.

3 Purification de ss-aIF1A

Le processus d'expression et de purification de ce facteur est très similaire à celui des protéines précédentes. Le vecteur utilisé (pET3alpa1Ass) a été cloné au laboratoire avant mon arrivée. Après une étape d'échange d'ion sur la colonne Q-Hiloal où la protéine n'est pas retenue, l'échantillon est appliqué sur la colonne S-Hiloal dans un tampon $A_1 10\% B_1$. Le gradient appliqué a la même pente que dans le cas de ss-aIF5B. Enfin, le tamis moléculaire est réalisé dans le tampon B_1 en raison de l'instabilité de la protéine observée à plus basse force ionique.

4 Clonage et purification du facteur pa-aIF5B

La protéine codée par le gène de pa-aIF5B contient une intéine de 394 résidus insérée après le 20^{ème} codon. La région C-terminale de la protéine a été amplifiée par PCR à partir de l'ADN chromosomique de *Pyrococcus abyssi*. Ce fragment, auquel ont été ajoutés un site *SmaI* correspondant aux résidus G19 et K20 d'une part et un site *SaII* en 3', a ensuite été cloné entre les sites *SmaI* et *XhoI* du vecteur pET3alpa produisant le vecteur pET3alpa-5BpaC. Le gène codant pour la protéine mature a été reconstitué par l'amplification puis le clonage du fragment correspondant aux résidus 1 à 19 entre les sites *NdeI* et *SmaI* du vecteur pET3alpa5BpaC permettant d'obtenir finalement le vecteur d'expression pET3alpa5Bpa.

La purification est identique à celle de ss-aIF5B excepté qu'elle ne comporte pas d'étape de précipitation au sulfate de streptomycine et que le gradient de la colonne S-Hiloal débute à 50mM NaCl. La protéine obtenue est homogène. Sa concentration est déduite du coefficient d'extinction molaire calculé de 0,631 cm²/mg. La protéine est concentrée à 10mg/mL. 8mg de protéine sont ainsi obtenus à partir d'un litre de culture.

5 Clonage et Purification de γ - Δ eIF5B

Cette protéine correspond au facteur eIF5B dont les 400 résidus N-terminaux sont supprimés. Le gène codant pour γ - Δ eIF5B a été amplifié par PCR à partir de l'ADN génomique de la souche W303a de *Saccharomyces cerevisiae* en introduisant un site *NdeI* et donc un codon de démarrage en place du résidu S395. Le fragment *NdeI-BamHI* est ensuite cloné dans un vecteur pET15blpa pour obtenir le plasmide pET15b5By.

La protéine produite à partir de ce plasmide comprend une étiquette His₆ en N-terminal. Le protocole d'expression est identique aux précédents. Les cellules sont ensuite lysées dans le tampon C₂ compatible avec l'application du surnageant sur 5mL de résine TALON. Suivant l'application, la résine est lavée par 50mL du tampon d'application A₂ avant l'élution de la protéine par ce même tampon complétement par 25mM d'imidazole. L'échantillon est alors dilué 5 fois dans le tampon A₂ avant d'être appliqué sur une colonne Q-Hiloal équilibrée dans un tampon A₂10%B₂. Un gradient de 80mL de 10% à 50% en tampon B₂ est ensuite appliqué. Les fractions contenant la protéine sont ensuite concentrées à un volume final de 3mL afin de permettre le dépôt sur une colonne Superdex75 équilibrée dans un tampon A₂50%B₂. La protéine élue sous un pic unique, sous forme homogène. La concentration en protéine est évaluée à l'aide du coefficient d'extinction molaire calculé de 0,588cm²/mg. La protéine est concentrée à 10mg/mL et 22mg de protéines ont pu être obtenus à partir d'1 litre de culture.

6 Purification des ARNt

L'ARNt_f^{Met}, l'ARNt₁^{Val}, l'ARNt^{Lys} ainsi que leurs variants ont été produits chez *E. coli* comme décrit précédemment dans la littérature ([Commans et al., 1995](#); [Guillon et al., 1992b](#); [Meinzel et al., 1991](#)). Le gène de l'ARNt initiateur de *Sulfolobus solfataricus* a été construit par l'assemblage d'oligonucléotides et clonage entre les sites *EcoRI* et *PstI* de pBSTNAV2 ([Meinzel et al., 1988](#)). Le gène cloné a ensuite été amplifié par PCR en ajoutant des sites *NdeI* et *XhoI* afin d'être transféré dans un vecteur pET3alpa pour l'exprimer dans une souche Rosetta(DE3)-pLacI/RARE. L'ARNt produit est ensuite purifié par une étape d'échange d'anions ([Meinzel et al., 1988](#)). L'ARNt purifié a ensuite été caractérisé par une analyse en SM. L'ARNt_f^{Met} de levure nous a été généreusement fourni par Gérard Keith (IGBMC, Strasbourg). Les capacités acceptrices en acide aminé de ces ARNt ont été mesurées entre 1000 et 1500pmol/A_{260nm}. Toutes les aminoacylations avec la méthionine marquée à l'isotope 35 du soufre (Amersham, diluée isotopiquement à 10000dpm/pmol) ont été réalisées par le mutant M547 de la méthionyl-ARNt synthétase d'*E. coli* ([Mellot et al., 1989](#)). La valine marquée au tritium (perkin Elmer, diluée isotopiquement à 5000dpm/mol) a été estérifiée par de la valyl-ARNt synthétase d'*E. coli* purifiée ([Brevet et al., 1989](#)). Enfin, la lysylation avec la lysine tritiée (amenée à 8000dpm/pmol) a été effectuée grâce à la lysyl-ARNt synthétase purifiée d'*E. coli* ([Commans et al., 1995](#)). Les ARNt ainsi aminoacylés sont précipités par l'éthanol en présence de 0,3M NaAc pH5,5 puis conservés à -20°C dans l'eau en petites quantités. Avant d'être utilisés, les niveaux d'aminocyclation des ARNt sont systématiquement vérifiés par précipitation par 5% d'acide trichloroacétique (TCA) puis filtration et comptage.

7 Tests de protection

Les réactions de protection des ARNt aminoacylés par les protéines e/alf5B ont été réalisées comme suit ([Yatime et al., 2004](#)). Des mélanges réactionnels de 120µL contiennent {20mM Hepes-NaOH (pH8,0); 100mM KCl; 500mM NaCl; 5mM MgCl₂; 1mM DTT; 0,1mM EDTA; 0,2mg/mL BSA (Roche); 5% glycérol; 1mM GTP et l'ARNt aminoacylé étudié à 80nM final}. Les protéines sont ajoutées à différentes concentration (de 1 à 120µM; le choix de la gamme est ajusté en fonction d'une première évaluation grossière de la constante de dissociation) dans les différents milieux réactionnels préparés de sorte à maintenir une concentration constante en sel qui influence la vitesse de déacylation. Les réactions sont ensuite incubées à 51°C, sauf indications contraires. Afin de déterminer les vitesses de déacylation, des échantillons de 20µL de mélange réactionnel sont prélevés à différents instants (de 5 à 30min.). Ces échantillons sont immédiatement précipités au TCA (5%) en présence d'ARN de levure pour faciliter le recueil de tous les ARNt sur des filtres GFC (Whatman). Les courbes de déacylation en fonction du temps sont ainsi obtenues. Dans tous les cas, les courbes ont pu être traitées telles des exponentielles. Les constantes de temps mesurées à différentes concentrations de protéine sont ensuite retranchées à celle d'un blanc sans protéine avant d'être tracées en fonction de la concentration en protéine. Les constantes de dissociation correspondantes sont déduites grâce à l'équation (8) du chapitre II après ajustement avec le logiciel MC-Fit ([Dardel, 1994](#)).

8 Cristallisation et cryoprotection

Les répétitions ont été réalisées en gouttes suspendues par la technique de diffusion de vapeur en mélangeant 2 μ L de liqueur mère (3,5mg/mL) et 2 μ L de solution de cristallisation pour former la goutte. Celle-ci est déposée sur une lamelle siliconée retournée sur un réservoir contenant 500 μ L de la condition de cristallisation. L'ensemble est scellé par de la graisse au silicone. Les criblages initiaux sont effectués à partir de matrices de solutions commerciales, d'une part, et de solutions préparées au laboratoire, d'autre part.

La meilleure condition permettant d'obtenir des cristaux d'aIF5B contient {0,1M Tris pH8,5 ; 0,25M sulfate d'ammonium et 31% de PEG3350}. Les cryoprotectants testés sont le PEG3350 (utilisé à 35% (v/v) final), le PEG400 (15%) et le glycérol (25%). Les solutions de cryoprotection sont obtenues par ajout de l'agent à la condition de cristallisation. Il a finalement été choisi le PEG3350, qui permet de conserver les cristaux et d'obtenir la meilleure diffraction.

III Le facteur eIF3

1 Clonages pour obtenir la bibliothèque de vecteurs

Les **tableaux 4.5 et 4.6** résument les vecteurs et les oligonucléotides utilisés dans les différents clonages. Les méthodes de clonage utilisées sont indiquées ; elles sont décrites plus précisément dans la partie I. La numérotation des résidus des oligonucléotides est celle des constructions pET15b1pa de départ. La description des différentes formes dans cette partie se réfère à la numérotation des protéines sauvages. Les séquences des protéines sauvages et dans les différentes constructions sont indiquées en Annexe 1.

2 Expression et purification

Une procédure standard est ici décrite. Les spécificités des différentes purifications sont détaillées plus loin. Pour produire les sous-unités isolées ou les complexes, des bactéries Rosetta(DE3)-pLacI/RARE sont transformées par électroporation en présence d'1 μ L du ou des plasmides. La transformation reprise dans 1mL de milieu riche non sélectif (2xTY) est laissée 1 heure à 37°C afin de laisser s'exprimer les différents gènes de résistance. 1L du milieu 2xTY complétement par les antibiotiques adéquats est inoculé puis mis en culture à 37°C. Après une nuit de culture, l'induction des bactéries à une DO_{650nm} d'environ 1,6 est réalisé par ajout d'IPTG à une concentration finale de 1mM et la culture est prolongée de 5 heures à 20°C.

Les cellules sont lysées par sonication dans 40mL de tampon C₂. Après centrifugation (30min. à 14500g), le surnageant est appliqué sur 5mL de résine TALON. Suivant l'application, la résine est lavée par 50mL du tampon d'application A₂. La protéine ou le complexe est élué par ce même tampon complétement par 125mM d'imidazole. L'hydrolyse de l'étiquette est réalisée au cours d'une dialyse destinée à extraire l'imidazole de l'échantillon. La quantité de protéine étiquetée récoltée de l'éluat est évaluée à l'aide du coefficient d'extinction molaire calculé à partir de la séquence en protéines. La thrombine est ajoutée à raison de 0,25UI par mg de protéine directement dans le boudin de dialyse. Le boudin est ensuite dialysé une nuit contre 1L de tampon T. Le dialysat est ensuite appliqué sur la résine TALON régénérée en milieu acide (pH5,0). La fraction de l'application qui passe au travers de la colonne est collectée. L'échantillon est alors dilué 5 fois dans le tampon A₂ avant d'être appliqué sur une colonne Q-Hiload (20mL de résine)

équilibrée dans un tampon A₂10%B₂. Un gradient de 200mL de 10% à 60% en tampon B₂ est ensuite appliqué. Lorsqu'un complexe contient la sous-unité eIF3a, l'élution est réalisée par un saut rapide de force ionique de 10 à 60% en tampon B₂ afin de contourner la forte affinité de cette sous-unité pour la résine. Les fractions contenant la protéine ou le complexe sont ensuite concentrées à un volume final de 3mL puis appliquées sur une colonne Superdex75 équilibrée dans un tampon A₂25%B₂.

L'application d'un échantillon sur la résine d'interaction hydrophobe nécessite d'amener la protéine à sa limite de solubilité par ajout de sulfate d'ammonium. La limite de solubilité est déterminée pour chaque protéine ou complexe étudié. L'application est alors réalisée sur la colonne (8mL de résine) équilibrée dans le tampon SA par ajouts successifs d'1 mL de l'échantillon puis 2 mL de tampon SA. Une fois l'échantillon appliqué, la colonne est lavée par 50mL de tampon SA. L'application d'un gradient de 50mL de A₂100%SA à 30% en tampon SA puis d'un gradient de 100mL de 30 à 0% en tampon SA permet d'éluer les protéines.

En fin de purification, la pureté est évaluée par SDS-PAGE et la concentration est déterminée par acquisition d'un spectre d'absorption. Enfin, l'échantillon est caractérisé par DLS. Cette dernière analyse est réalisée sur 12μL d'un échantillon à une concentration de l'ordre du mg/mL, centrifugé pendant 15 min. à 14500g afin d'éviter que la mesure soit perturbée par des poussières. Chaque mesure est répétée 3 fois et la qualité du traitement est estimée par la forme de la fonction d'autocorrélation.

3 Les sous-unités isolées

a eIF3aN1, eIF3aN2 (pET15bXSB-eIF3aN1, pET15bXSB-eIF3aN2)

La purification de cette sous-unité est réalisée en trois étapes : TALON, Q-Hiload et tamisage moléculaire. L'analyse SDS-PAGE révèle une forte protéolyse de la sous-unité. 5mg de protéine sont obtenus à partir d'1 L. de culture.

b eIF3bC1 (pET32b-eIF3bC1)

La purification est réalisée par les étapes : TALON, Superdex puis Q-Hiload. L'application sur tamis moléculaire révèle la présence de deux pics dont les différences de temps de rétention pourraient indiquer une dimérisation de la protéine. Celle-ci a pu être confirmée en injectant à nouveau le sommet d'un des pics du tamis moléculaire sur cette même colonne créant un profil comportant également deux pics. L'équilibre entre les deux formes de la protéine a pu être clairement mis en évidence par l'action d'un tensioactif dissociant les complexes, le Triton X-100. L'ajout à cette même fraction du tamis moléculaire 0,1% v/v de Triton X-100 permet de n'observer qu'un pic unique, identifié comme celui de l'espèce monomérique. L'utilisation de cette quantité de Triton X-100 étant *a priori* compatible avec un maintien de la structure de la protéine. La forme monomérique ainsi que les différents pics obtenus en l'absence de tensioactif ont été testés pour la cristallisation.

L'excision de l'étiquette de la protéine ne modifie pas le profil d'élution du tamis moléculaire, indiquant que la dimérisation intervient par un domaine interne à eIF3bC1.

c eIF3gC1 (pET15blpa-eIF3gC1)

La protéine possédant un motif RRM dans sa séquence et le spectre d'absorption de la protéine présentant effectivement un pic à 280nm mais une absorption anormalement élevée à 260nm, nous avons pensé à une interaction pontée par des acides nucléiques. Suivant l'étape d'accrochage sur TALON, une étape d'échange d'ion à plus haute force ionique (500mM) permet de recueillir la protéine dans la fraction qui passe au travers de la colonne alors que les acides nucléiques sont retenus. L'application sur le tamis permet d'obtenir un pic symétrique correspondant à eIF3gC1 avec un bon facteur de pureté. Le spectre d'absorption du produit purifié ne semble plus montrer la présence d'acides nucléiques.

d eIF3i (pET15blpa-eIF3i)

La purification de cette sous-unité comporte les étapes TALON, retrait de l'étiquette, Q-Hiload et Superdex. Afin d'obtenir la cristallisation, le premier pic de la colonne d'échange d'ions doit être sélectionné puis tamisé.

4 Purifications de complexes à deux partenaires

a eIF3bC2-3cC1 (pET32b-eIF3bC2 + pCDF-3cC1)

Suivant l'étape TALON, les deux protéines coéluent lors de l'étape d'échange d'ions, le profil d'un tamis moléculaire réalisé à une force ionique semblable à celle à laquelle les protéines sont éluées de la colonne précédente présente clairement deux pics attribuables à chacune des protéines. Un pic de cette colonne précède directement celui d'eIF3bC2, le gel SDS-PAGE révèle une bande diffuse très étalée dans la région des poids moléculaires compris entre 30 et 76kDa ainsi qu'une bande assez faible co-migrant avec eIF3bC2. Ce pic pourrait correspondre à un dimère entre des fragments protéolysés d'eIF3cC1 et eIF3bC2. De plus, cette bande diffuse semble contaminer le pic correspondant à eIF3bC2. Clairement, s'il y a interaction celle-ci doit être faible. Ce résultat est cohérent avec le fait que peu de résidus impliqués dans cette liaison sont maintenus au sein d'eIF3bC2.

b eIF3i-3gC1 (pCOLA-3i-3gC1)

La purification comporte une étape TALON suivie d'une étape d'échange d'ions et d'un tamisage moléculaire. Deux pics sont résolus sur le profil d'éluion de la résine échangeuse d'ion. Aucune différence de composition des fractions correspondantes n'est détectable par une analyse en SDS-PAGE. Les échantillons fournis pour une identification en SM n'ont pas pu être suffisamment dessalés pour une analyse en masse totale en conditions natives. Les analyses de SM en conditions dénaturantes donnent des résultats très similaires pour les deux pics excepté la présence de trois pics non attribuables dans le spectre du deuxième pic.

c eIF3a-3cC1 (pCDF-3a-3cC1)

L'étape TALON est suivie par une étape d'échange d'ions (éluion par saut de force ionique) et un tamisage moléculaire.

d eIF3bC1-3i (pET32b-3bC1 + pCOLA-3i)

Suivant les étapes TALON, d'échange d'ion et de tamisage. Deux pics sont résolus lors du passage sur tamis moléculaire. Les temps de rétention observés sont compatibles avec une dimérisation. L'application séparée de chacun des pics génère à nouveau deux pics sur une colonne échangeuse d'ion.

e eIF3bC1-3a (pET32b-3bC1 + pCDF-3a)

Le profil d'élution de la résine TALON ne permet pas de détecter eIF3a que ce soit dans l'élution, le surnageant ou encore le culot de sonication.

f eIF3cC1-3i (pET15blpa-3i + pCDF-3cC1)

La protéine eIF3cC1 n'apparaît cependant pas dans l'élution de la résine IMAC alors qu'elle semble être présente dans la fraction qui passe au travers de la colonne, indiquant l'absence de formation de dimère.

5 Purifications de complexes trimériques

a eIF3bC1-3i-3gC1 (pET32b-3bC1 + pCOLA-3i-3gC1)

Cela a permis de mettre en évidence la présence de trois pics distincts lors du passage sur tamis moléculaire après les étapes TALON et d'échange d'ions. Le dernier pic correspond à l'excès de la sous-unité eIF3bC1. Le second pic élué à un temps de rétention comparable à celui du complexe isolé précédemment alors que le premier pic, d'une hauteur comparable, pourrait correspondre à la formation d'un dimère (3bC1-i-gC1)₂. Il s'avère que l'application sur tamis moléculaire d'un des pics du tamis moléculaire se traduit par l'élution de ce même pic, ce qui diffère notablement de la situation d'eIF3bC1. Il a pu être vérifié au cours de cette purification que l'isolement des pics issus du tamis moléculaire est stable même après concentration en vue de réaliser des tests de cristallisation. L'application d'une fraction du concentrat ne génère qu'un pic unique sur tamis moléculaire.

b eIF3a/3aN1/3aN2-3bC1-3cC1 (pET32b-3bC1 + pCDF-3a/3aN1/3aN2-3cC1)

Des purifications comportant la séquence complète d'étapes de purifications ou en omettant les étapes de retraits d'étiquette ou de chromatographie d'interaction hydrophobe (limite de solubilité 0,7M SA) n'ont pas permis d'isoler un complexe homogène.

c eIF3cC1-3i-3gC1 (pET15blpa-eIF3cC1/pCDF-3cC1 + pCOLA-3i-3gC1)

Une faible quantité de protéine est éluée de la résine TALON avec le dimère eIF3i-3gC1 pour les deux constructions permettant de produire eIF3cC1 (avec ou sans étiquette). Les quantités de cette sous-unités sont cependant faibles et le passage sur la résine échangeuse d'ion bien qu'indiquant une co-élution des protéines ne permet pas de visualiser une stœchiométrie convenable sur SDS-PAGE.

6 Purifications de complexes tétramériques

a eIF3a-3cC1-3i-3gC1 (pCDF-3a-3cC1 + pCOLA-3i-3gC1)

La purification est réalisée en trois étapes : TALON, tamisage moléculaire et échange d'ion.

b eIF3bC1-cC1-i-gC1

La purification de ce complexe comporte une étape d'accrochage sur résine TALON, de retrait d'étiquette, d'échange d'ion, de tamisage moléculaire et enfin d'interaction hydrophobe (limite de solubilité 0,7M SA).

7 Purifications du complexe pentamérique

pET32-3bC1 + pCDF-3a/3aN1/3aN2-3cC1 + pCOLA-3i-3gC1

ou

pET15bXSB-3a/3aN1/3aN2 + pCDF-3bC1-3cC1 + pCOLA-3i-3gC1

La séquence comporte généralement un accrochage spécifique sur résine TALON puis plusieurs polissages sur tamis moléculaire. L'étape d'échange d'ions par saut de force ionique a été ajoutée mais semble défavoriser une stoechiométrie correcte.

Le cocassage réalisé en mélangeant le culot de bactéries correspondant à 1/3L de culture surproduisant eIF3a (pET15bXSB-3aN1) avec le culot d'1 L. de culture surproduisant l'hétérotétramère eIF3bC1-cC1-i-gC1 (pCDF-3bC1-3cC1 + pCOLA-3i-3gC1) n'a pas permis d'obtenir une stoechiométrie satisfaisante après trois application sur tamis moléculaire en raison de l'excès d'eIF3aN1.

IV La sous-unité eIF3i, cristallisation

La séquence des oligonucléotides utilisés dans les différents clonages sont reportés dans le **tableau 4.7**.

1 Cristallisation

La protéine a été cristallisée manuellement par la méthode de diffusion de vapeur en goutte assise dans de boîtes à 96 puits. La solution protéique à 5mg/mL dans un tampon {10mM MOPS pH6,7 ; 3mM 2-mercaptoéthanol ; 250 mM NaCl} a été utilisée. Les gouttes sont réalisées en apposant délicatement 1µL de solution de protéine sur 1µL de la solution de cristallisation (1,8M Na/KPO₄ pH 8,2). Les cristaux sont obtenus à 4°C après 3 semaines.

Pour mener les expériences de diffraction, les cristaux ont été trempés dans une solution composée de la solution de cristallisation additionnée de 20% (v/v) glycérol.

2 Expression et purification des protéines sélénométhionylées

La souche d'*E. coli* B834(DE3) auxotrophe pour la méthionine est électrotransformée par 1 μ L de plasmide. Les cellules sont mises en culture dans 50mL de milieu riche 2xTY. La préculture est incubée à 37°C pendant une nuit, la quantité de cellules est estimée par mesure de densité optique à 650nm. Les cellules sont récupérées par centrifugation (8000g, 4°C, 20 min.). Le culot de bactéries est rincé deux fois avec 50mL de milieu minimal M9 puis repris dans du milieu minimal M9 afin d'atteindre une densité optique de 100 DO_{650nm}. 1L de milieu M9 contenant du glucose 0,2%, l'antibiotique permettant de sélectionner le plasmide et 40mg de chacun des acides aminés est inoculé par 1 mL de la préculture lavée (inoculation à 0,1 DO_{650nm}). 15mg de sélénométhionine (SIGMA) sont finalement ajoutés à la culture. Celle-ci est incubée à 37°C jusqu'à ce que la DO_{650nm} atteigne 0,8-1,0. La surproduction de la protéine est alors induite par ajout d'IPTG à une concentration finale de 0,1mM et l'induction est poursuivie à 20°C pendant 5 heures.

Alternativement, la surproduction a été testée en milieu autoinductible ([Studier, 2005](#)). Ce milieu minimal contient du lactose et du glucose, tous les acides aminés et de la sélénométhionine. La méthionine introduite permet de limiter l'effet toxique de la sélénométhionine. De même que précédemment, la culture est inoculée à 0,1DO_{650nm} à partir d'une préculture lavée. Lorsque les cellules auront consommé le glucose du milieu, la présence de lactose en l'absence de glucose permet d'induire la culture. La culture est incubée une journée à 37°C puis transférées à 20°C jusqu'à atteindre une concentration d'environ 4-5DO_{650nm}/mL.

Les protéines séléméthionylées ainsi produites et comportant une étiquette sont purifiées suivant le protocole de la protéine native. Dans le cas d'un étiquetage N-terminal, l'impossibilité du retrait de l'étiquette a conduit à mener les étapes suivantes de purification en la présence de celle-ci.

3 Production d'eIF3i sous forme native

Le plasmide pET3alpa-eIF3i a été obtenu par sous-clonage du vecteur pET15blpa-eIF3i hydrolysé par les enzymes *NdeI* et *CelI* puis ligature de l'insert dans un vecteur pET3alpa hydrolysé par les mêmes enzymes.

La protéine sélénométhionylée n'a pas pu être identifiée après une étape d'échange d'ions.

4 Production d'eIF3i fusionnée à une étiquette C-terminale

Le clonage dans un vecteur de type pET32b a été réalisé à l'aide des enzymes de restriction *NdeI* et *XhoI*, à partir du plasmide pET15plpa-eIF3i.

Le protocole de purification est identique à celui utilisé pour la protéine fusionnée en N-terminal.

5 Production des mutants de délétion des résidus C-terminaux

La modification du plasmide pET15blpa-eIF3i a été réalisée par mutagenèse dirigée en introduisant un codon STOP au niveau du résidu Glu341 (pET15blpa-3iSTOP1) et au niveau du résidu Ala336 (pET15blpa-3iSTOP2).

La modification du plasmide pET32b-eIF3i a été réalisée par mutagenèse de délétion. La forme eIF3iDEL1 est dépourvue des résidus Ala337 à Asn347 et la forme eIF3iDEL2 ne comporte pas la séquence C-terminale Gln344 à Asn347.

6 Obtention des mutants cystéines multiples

La méthode QuickChange MultiSite Directed Mutagenesis (Stratagene) a été utilisée suivant les recommandations du protocole fourni. Dans ce cas, un oligonucléotide aller correspondant à chacune des mutations est ajouté à la matrice. L'ajout d'un mélange d'enzymes contenant une ADN-polymérase et une ADN-ligase permet de produire un ADN simple brin comportant les différentes mutations. Une digestion par l'enzyme de restriction *DpnI* permet d'hydrolyser les plasmides et les brins parentaux. La suite de la procédure est la même que pour une mutagenèse ponctuelle.

Deux quadruples mutants ont ainsi été obtenus : eIF3iC62AC86AC104SC238A (pET15blpa-3iMulti3) et eIF3iC67AC86AC104SC238A (pET15blpa-3iMulti9).

Plasmide synthétisé		
pET32b-eIF3i	Oligonucléotide aller	5'-CGCGGCAGCCATATGAAGGCT-3'
	Oligonucléotide retour	5'-CCCCTCGAGAGAACC GCGTGGCACCAGATTAGCTTCTTGCAT-3'
pET15blpa-3iSTOP1	Oligonucléotide aller	5'-GCCGCCGAAGCTAAATGACACATGCAAGAAGCT-3'
	Oligonucléotide retour	5'-AGCTTCTTGCATGTGTCAATTTAGCTTCGGCGGC-3'
pET15blpa-3iSTOP2	Oligonucléotide aller	5'-TACGACGTGGAGAAGTGAGCCGAAGCTAAAGAG-3'
	Oligonucléotide retour	5'-CTCTTTAGCTTCGGCTCACTTCTCCACGTCGTA-3'
pET32b-3iDEL1	Oligonucléotide aller	5'-AAATACGACGTGGAGAAGGCCCTGGTGCCACGCGTTCTCTC-3'
	Oligonucléotide retour	5'-GAGAGAACCGCGTGGCACCAGGGCCTTCTCCACGTCGTATTT-3'
pET32b-3iDEL2	Oligonucléotide aller	5'-GCCGAAGCTAAAGAGCACATGCTGGTGCCACGCGTTCTCTC-3'
	Oligonucléotide retour	5'-GAGAGAACCGCGTGGCACCAGCATGTGCTCTTTAGCTTCGGC-3'
pET15blpa-3iMulti3	IScC82A aller	5'-TCTATCGACGTCGACGCTTTCACCAAGTACTGT-3'
	IScC106A aller	5'-GTGTCAAACGGGCAAGCTGTTGCTACATGGAAA-3'
	IScC124S aller	5'-GTTGAATTTTCCCATCCGGTAACTACTTTCAT-3'
	IScC258A aller	5'-AAATACGAAACAGACGCTCCTCTAAACACCGCC-3'
	IScC87A aller	5'-TGTTTCACCAAGTACGCTGTTACAGGAAGTGCT-3'
pET15blpa-3iMulti9	IScC106A aller	5'-GTGTCAAACGGGCAAGCTGTTGCTACATGGAAA-3'
	IScC124S aller	5'-GTTGAATTTTCCCATCCGGTAACTACTTTCAT-3'
	IScC258A aller	5'-AAATACGAAACAGACGCTCCTCTAAACACCGCC-3'

Tableau 4.7 : Oligonucléotides pour produire les différents plasmides permettant de produire eIF3i.

Ori./Rés.	MCS	Protéine	Plasmide obtenu	Plasmide matrice	Plasmide de destination	Technique	Oligonucléotide aller	Site 5'	Oligonucléotide retour	Site 3'	Contre-sélection
ColE1 / AmpR		eIF3aN1	pET15blpa-3aN1	pET15blpa-eIF3a		Mut. DEL.	MUTAC1-15F		MUTAC1-15R		<i>DpnI</i>
		eIF3aN1(nt)	pET15blpa-3aN1(nt)	pET15blpa-3aN1		Mut. DEL.	3A-TAG-F		3A-TAG-R		<i>DpnI</i>
		eIF3aN2	pET15blpa-3aN2	pET15blpa-eIF3a		Mut. DEL.	MUTAC2-15F		MUTAC2-15R		<i>DpnI</i>
		eIF3aN2(nt)	pET15blpa-3aN2(nt)	pET15blpa-3aN2		Mut. DEL.	3A-TAG-F		3A-TAG-R		<i>DpnI</i>
		eIF3b	pET32b-3b	pET15blpa-eIF3b	pET32b	Clonage ER	Prt1- <i>NdeI</i>	<i>NdeI</i>	3B-763-XO-R	<i>XhoI</i>	<i>HindIII</i>
		eIF3bC1	pET32b-3bC1	pET15blpa-eIF3b	pET32b	Clonage ER	3B-193-NE-F	<i>NdeI</i>	3B-763-XO-R	<i>XhoI</i>	<i>HindIII</i>
		eIF3bC1thr	pET32b-3bC1thr	pET32b-3bC1		Mut. INS.	MUTA3BC1THROMF2		MUTA3BC1THROMR2		<i>DpnI</i>
		eIF3bC2	pET32b-3bC2	pET15blpa-eIF3b	pET32b	Clonage ER	3B-589-NE-F	<i>NdeI</i>	3B-763-XO-R	<i>XhoI</i>	<i>HindIII</i>
		eIF3bN1	pET32b-3bN1	pET15blpa-eIF3b	pET32b	Clonage ER	Prt1- <i>NdeI</i>	<i>NdeI</i>	3B-175-XO-R	<i>XhoI</i>	<i>HindIII</i>
		eIF3cC1	pET15blpa-3cC1	pET15blpa-eIF3c		Mut. DEL.	MUT3C15BF		MUT3C15BR		<i>DpnI</i>
	eIF3gC1	pET15blpa-3gC1	pET15blpa-eIF3g		Mut. DEL.	MUT3GC1-15BF		MUT3GC1-15BR		<i>DpnI</i>	
CloDF13 / SmR	1	eIF3a	pCDF-3a	pET15bXSB-eIF3a	pCDFDuet-1	Clonage ER	3A1BAMF2	<i>BamHI</i>	3A966NOTR2	<i>NotI</i>	<i>XhoI</i>
	1	eIF3aN1	pCDF-3aN1-3cC1	pCDF-3a-3cC1		Mut. DEL.	MUTAC1-CDF		MUTAC1-CDR		<i>DpnI</i>
	1	eIF3aN1	pCDF-3aN1	pCDF-3a		Mut. DEL.	MUTAC1-CDF		MUTAC1-CDR		<i>DpnI</i>
	1	eIF3aN2	pCDF-3aN2-3cC1	pCDF-3a-3cC1		Mut. DEL.	MUTAC2-CDF		MUTAC2-CDR		<i>DpnI</i>
	1	eIF3aN2	pCDF-3aN2	pCDF-3a		Mut. DEL.	MUTAC2-CDF		MUTAC2-CDR		<i>DpnI</i>
	1	eIF3bC1thr	pCDF-3bC1-3cC1	pET32b-3bC1thr	pCDF-3cC1	Clonage ER	3BC1TF- <i>NcoI</i>	<i>NcoI</i>	3BC1TR-NOTI	<i>NotI</i>	<i>PstI</i>
	1	eIF3bC1(nt)	pCDF-3bC1(nt)-3cC1	pCDF-3bC1-3cC1		Mut. DEL.	3BC1-TAG-F		3BC1-TAG-R		<i>DpnI</i>
	2	eIF3c	pCDF-3c	pET15blpa-eIF3c	pCDFDuet-1	Clonage ER	3C-1-KPN-F	<i>KpnI</i>	3C-812-XHOR	<i>XhoI</i>	
	2	eIF3c	pCDF-3a-3c	pET15blpa-eIF3c	pCDF-3a	Clonage ER	3C-1-KPN-F	<i>KpnI</i>	3C-812-XHOR	<i>XhoI</i>	
	2	eIF3cN1	pCDF-3cN1	pET15blpa-eIF3c	pCDFDuet-1	Clonage ER	3C-1-KPN-F	<i>KpnI</i>	3C-206-SAL-R	<i>Sall/XhoI</i>	<i>XhoI</i>
	2	eIF3cC1	pCDF-3a-3cC1	pCDF-3a-3c		Mut. DEL.	MUTAC1F-KPN1		MUTAC1R		<i>KpnI/DpnI</i>
	2	eIF3cC1	pCDF-3cC1	pCDF-3c		Mut. DEL.	MUTAC1F-KPN1		MUTAC1R		<i>KpnI/DpnI</i>
	2	eIF3cC1(t)	pCDF-3bC1-3cC1(t)	pCDF-3bC1-3cC1		Mut. INS.	3CC1-TAG-F		3CC1-TAG-R		<i>DpnI</i>
2	eIF3cC1(t)	pCDF-3bC1(nt)-3cC1(t)	pCDF-3bC1-3cC1		Mut. INS.	3CC1-TAG-F		3CC1-TAG-R		<i>DpnI</i>	
2	eIF3cC1(tt)		pCDF-3bC1(nt)-3cC1(t)		Mut. INS.	3CC1-THROM-F		3CC1-THROM-R		<i>DpnI</i>	
ColA / KanR	1	eIF3i	pCOLA-3i	pET15blpa-eIF3i	pCOLADuet	Clonage ER	3I-1-BAM-F	<i>BamHI</i>	3I-347-NOT-R	<i>NotI</i>	<i>PstI</i>
	1	eIF3i(nt)	pCOLA-3i(nt)-3gC1	pCOLA-3i-3gC1		Mut. DEL.	3I-TAG-F		3I-TAG-R		<i>DpnI</i>
	2	eIF3g	pCOLA-3i-3g	pET15b-eIF3g	pCOLA-3i	Clonage ER	3G1MFEF	<i>MfeI</i>	3G378XHORBIS	<i>XhoI</i>	<i>KpnI</i>
	2	eIF3gC1	pCOLA-3i-3gC1	pCOLA-3i-3g		Mut. DEL.	MUTA3G-ALLER		MUTA3G-RETOUR		<i>DpnI</i>

Tableau 4.5 : Bibliothèque de vecteurs. Récapitulatif des clonages et mutagenèses. Clonage ER. : Clonage par enzymes de restrictions. Mut. DEL. : Mutagenèse de délétion. Mut. INS. : Mutagenèse d'insertion

Oligonucléotide aller	Séquence (5'-3')	Oligonucléotide retour	Séquence (5'-3')
MUTAC1-15F	GCCATCGAACAGAGACTAGCCTAACAGACAACGCAAAGGAT	MUTAC1-15R	ATCCTTTTGC GTTGTCTGTTAGGCTAGTCTCTGTTTCGATGGC
3A-TAG-F	ACTTTAAGAAGGAGATATACCATGCTCGAGGCCCCCCACCA	3A-TAG-R	TGGTGGGGGGGCTCGAGCATGGTATATCTCCTTCTTAAAGT
MUTAC2-15F	GAAATCGAAGAAGCTGCTGCCTAACAGACAACGCAAAGGAT	MUTAC2-15R	ATCCTTTTGC GTTGTCTGTTAGGCAGCAGCTTCTTCGATTTTC
3A-TAG-F	ACTTTAAGAAGGAGATATACCATGCTCGAGGCCCCCCACCA	3A-TAG-R	TGGTGGGGGGGCTCGAGCATGGTATATCTCCTTCTTAAAGT
Prt1-NdeI		3B-763-XO-R	AGTTCATGACTCGAGTTCGACCTTTTCCTTTGTTTTC
3B-193-NE-F	AAATCCTGGCATATGGACGACAAAGTGAGAGATCAG	3B-763-XO-R	AGTTCATGACTCGAGTTCGACCTTTTCCTTTGTTTTC
MUTA3BC1THROMF	AAGGAAAAGGTGCAACTCCTGGTGCCACGCGTTCTGAGCACCA	MUTA3BC1THROMR2	GTGGTGGTGGTGGTGTCTCAGAACCGCGTGGCACCAGGAGTTTCG
3B-589-NE-F	GAAAAGAATCATATGGACAATAATGATGTTTCAGCT	3B-763-XO-R	AGTTCATGACTCGAGTTCGACCTTTTCCTTTGTTTTC
Prt1-NdeI		3B-175-XO-R	CATATCAGGCTCGAGGAATTCCGGTGTCAAAGTCGTC
MUT3C15BF	GTGCCGCGCGGCAGCCATATGCAACAAAAGTGGGGGACTCCA	MUT3C15BR	TGGAGTCCCCAGTTTTGTTGCATATGGCTGCCGCGCGGCAC
MUT3GC1-15BF	CTGGTGCCGCGCGGCAGCCATATGGAAAAGTCCATAAATCTGT	MUT3GC1-15BR	AACAGATTTATGGACTTTTTCCATATGGCTGCCGCGCGGCACCA
3A1BAMF2	GCCGCGCGGGGATCCTATGCTTGAGGCCCCCCACC	3A966NOTR2	TCCTTTTGC GCGCGCCGCTTACCTGCCCCCTTGCCCAA
MUTAC1-CDF	GCCATCGAACAGAGACTAGCCTAAGCGGCCGCATAATGCTTA	MUTAC1-CDR	TAAGCATTATGCGGCCGCTTAGGCTAGTCTCTGTTTCGATGGC
MUTAC1-CDF	GCCATCGAACAGAGACTAGCCTAAGCGGCCGCATAATGCTTA	MUTAC1-CDR	TAAGCATTATGCGGCCGCTTAGGCTAGTCTCTGTTTCGATGGC
MUTAC2-CDF	GAAATCGAAGAAGCTGCTGCCTAAGCGGCCGCATAATGCTTA	MUTAC2-CDR	TAAGCATTATGCGGCCGCTTAGGCAGCAGCTTCTTCGATTTTC
MUTAC2-CDF	GAAATCGAAGAAGCTGCTGCCTAAGCGGCCGCATAATGCTTA	MUTAC2-CDR	TAAGCATTATGCGGCCGCTTAGGCAGCAGCTTCTTCGATTTTC
3BC1TF-NcoI	TAAGAAGGAGATATACCCATGGACGACAAAAGTGAGA	3BC1TR-NotI	CGGGCTTTGTTAGCAGCGGCCGCTCAGTGGTGGTGGTGGTGG
3BC1-TAG-F	ACAAAGGAAAAGGTGCAACTCTGAGCGGCCGCATAATGCTTA	3BC1-TAG-R	TAAGCATTATGCGGCCGCTCAGAGTTCGACCTTTTCCTTTGT
3C-1-KPN-F	CCGCGCGGCGGTACCATGCCCGGGCTTGAGTCCCGT	3C-812-XHOR	CGCGGATTTCTCGAGTCAACGACGATTTGATGGTGG
3C-1-KPN-F	CCGCGCGGCGGTACCATGCCCGGGCTTGAGTCCCGT	3C-812-XHOR	CGCGGATTTCTCGAGTCAACGACGATTTGATGGTGG
3C-1-KPN-F	CCGCGCGGCGGTACCATGCCCGGGCTTGAGTCCCGT	3C-206-SAL-R	TGAGTCTTCGTCGACTTAAGAACTTTCTTGACTCTTTG
MUTAC1F-KPN1	ATCGCTGACGTCGCTACCATGCAACAAAAGTGGGGGACTCCA	MUTAC1R	TGGAGTCCCCAGTTTTGTTGCATGGTAGCGACGTCAGCGAT
MUTAC1F-KPN1	ATCGCTGACGTCGCTACCATGCAACAAAAGTGGGGGACTCCA	MUTAC1R	TGGAGTCCCCAGTTTTGTTGCATGGTAGCGACGTCAGCGAT
3CC1-TAG-F	TAAGAAGGAGATATACATATGGGCAGCAGCCATCATCATCATCAT	3CC1-TAG-R	TATCCAATTGAGATCTGCGTGATGATGATGATGATGGCTGCTGC
3CC1-TAG-F	TAAGAAGGAGATATACATATGGGCAGCAGCCATCATCATCATCAT	3CC1-TAG-R	TATCCAATTGAGATCTGCGTGATGATGATGATGATGGCTGCTGC
3CC1-THROM-F	CATCATCATCATCATCACAGCAGCGCCTGGTGCCGCGCGGCAG	3CC1-THROM-R	TATCCAATTGAGATCTGCGTGATGGCTGCCGCGCGGCACCAGGCCG
3I-1-BAM-F	GCCGCGCGGGGATCCTATGAAGGCTATCAAATTAAC	3I-347-NOT-R	GAGCATCAAGCGGCCGCTTAATTAGCTTCTTGATGTG
3I-TAG-F	ACTTTAATAAGGAGATATACCATGAAGGCTATCAAATTAACA	3I-TAG-R	TGTTAATTTGATAGCCTTCATGGTATATCTCCTTATTAAGT
3G1MFEF	TATGCCCGGCAATTGGAGTGAAGTTGCACCAGAGGA	3G378XORBIS	GCGGATTCTCTCGAGCTATTCCTTAACCTTAGGTTT
MUTA3G-ALLER	GTATAAGAAGGAGATATACATATGGAAAAGTCCATAAATCTGTT	MUTA3G-RETOUR	AACAGATTTATGGACTTTTTCCATATGTATATCTCCTTCTTATAC

Tableau 4.6: Oligonucléotides utilisés pour obtenir la bibliothèque de vecteurs. L'ordre des oligonucléotides est le même que dans le tableau précédent.

Annexes

Séquences d'eIF3a

eIF3a native (MW = 110kDa; $\epsilon = 0,635\text{cm}^2/\text{mg}$; pI = 6,54)

10	20	30	40	50	60	70
MAPPPFRPEN	AIKRADELIS	VGEKQAALQS	LHDFITARRI	RWATPSTVEP	VVFKFLEIGV	ELKKGKLLKD
80	90	100	110	120	130	140
GLHQYKKLIQ	GSTEGLVSVG	AVARKFIDL	ESKIASEQTR	ADELQKQEID	DDLEGGVTPE	NLLISVYESD
150	160	170	180	190	200	210
QSVAGFNDEA	ITSWLRF TWE	SYRAVL DL LR	NNALLEITYS	G V VK K TMHFC	LKYQRKNEFK	RLAEMLRQHL
220	230	240	250	260	270	280
DAANYQQSKS	GNNLV DL SDA	DTLQRY LD QR	FQQVDVSVKL	ELWHEAYRSI	EDVFHLMKIS	KRAPK P STLA
290	300	310	320	330	340	350
NYYENLVKVF	FVSGD PL LHT	TAWKKFYKLY	STNPRATEEE	FKTYSSTIFL	SAISTQ L DEI	PSIGYD P HLR
360	370	380	390	400	410	420
MYRLLNLDAK	PTRKEM L QSI	IEDESIYGKV	DEELKELYDI	IEVNF D VDTV	KQQL EN LLVK	LSSKTYFSQY
430	440	450	460	470	480	490
IAPLRDVIMR	RVFVAASQKF	TTVSQSELYK	LATLPAPLDL	SAWDIEKSL	QAAVEDYVSI	TIDHESAKVT
500	510	520	530	540	550	560
FAKDPFDIFA	STASKEVSEE	ENTEPEVQEE	KEETDEALGP	QETEDGEEKE	EESDPVIIRN	SYIHNKLEL
570	580	590	600	610	620	630
SNVLHDVDSF	NNASYMEKVR	IARETLIKKN	KDDLEKISKI	VDERVKRSQE	QKQKHMEHAA	LHAEQDAEVR
640	650	660	670	680	690	700
QQRILEEKAA	IEAKLEEEAH	RRLIEKKKRE	FEAIKEREIT	KMITEVNAKG	HVYIDPNEAK	SLDLDTIKQV
710	720	730	740	750	760	770
IIAEVSKNKS	ELESRMEYAM	KKLDHTERAL	RKVELPLLQK	EVDKLQETDT	ANYEAMKKKI	VDAAKAEYEA
780	790	800	810	820	830	840
RMADRKNLVM	VYDDYLK F KE	HVSGTKESEL	AAIRNQK K AE	LEAAKKARIE	EVRKRRYEEA	IARRKEEIAN
850	860	870	880	890	900	910
AERQKRAQEL	AEATRQREI	EEAAAKSTP	YSFRAGNREP	PSTPSTLPKA	TVSPDKAKLD	MIAQKQREME
920	930	940	950	960		
EAIEQRLAGR	TAGGSSPATP	ATPATPATPT	PSSGPKKMTM	AEKLRAKRLA	KGGR	

eIF3a-His6 (MW = 113kDa ; $\epsilon = 0,622\text{cm}^2/\text{mg}$; pI = 6,81) pET15b~~l~~pa-eIF3a

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLV PRGSH	MLEAPPPFRP	ENAIKRADEL	ISVGEKQAAL	QSLHDFITAR	RIRWATPSTV
80	90	100	110	120	130	140
EPVVFKFLEI	GVELKKGKLL	KDGLHQYKKL	IQGSTEGLV	VGAVARKFID	LVESKIASEQ	TRADELQKQE
150	160	170	180	190	200	210
IDDDLEGGVT	PENLLISVYE	SDQSVAGFND	EAITSWLRFT	WESYRAVL DL	LRNNALLEIT	YSGVVK T MH
220	230	240	250	260	270	280
FCLKYQRKNE	FKRLAEMLRQ	HLDAANYQQS	KSGNNLV DL S	DADTLQRYLD	QRFQQVDVSV	KLELWHEAYR
290	300	310	320	330	340	350
SIEDVFHLMK	ISKRAPK P ST	LANYYENLVK	VFFVSGD PL L	HTTAWKKFYK	LYSTNPRATE	EEFKTYSSTI
360	370	380	390	400	410	420

Annexe 1 : Séquence de la sous-unité eIF3a produite à partir des différentes constructions indiquées en italique. La masse molaire (MW), le coefficient d'extinction molaire(E) et le point isolélectrique (pI) ont été calculés à partir des séquences dans DNAid ([Dardel and Bensoussan, 1988](#)). Les séquences apparaissant surlignées en gris sont excisées par l'action de la thrombine.

FLSAISTQLD	EIPSIGYDPH	LRMYRLLNLD	AKPTRKEMLQ	SIIEDESIYG	KVDEELKELY	DIIEVNFVDV
430	440	450	460	470	480	490
TVKQQLLENLL	VKLSSKTYFS	QYIAPLRDVI	MRRVFVAASQ	KFTTVSQSEL	YKLATLPAPL	DLSAWDIEKS
500	510	520	530	540	550	560
LLQAAVEDYV	SITIDHESAK	VTFADKPPFDI	FASTASKEVS	EEENTEPEVQ	EEKEETDEAL	GPQETEDGEE
570	580	590	600	610	620	630
KEEESDPVII	RNSYIHNKLL	ELSNVLHDVD	SFNNASYMEK	VRIARETLIK	KNKDDLEKIS	KIVDERVKRS
640	650	660	670	680	690	700
QEQQKQKHEH	AALHAEQDAE	VRQQRILEEK	AAIEAKLEEE	AHRRLIEKKK	REFEAIKERE	ITKMITEVNA
710	720	730	740	750	760	770
KGHVYIDPNE	AKSLDLDTIK	QVIIAEVSKN	KSELESMEY	AMKKLDHTE	ALRKVELPLL	QKEVDKLQET
780	790	800	810	820	830	840
DTANYEAMKK	KIVDAAKAEY	EARMADRKNL	VMVYDDYLKF	KEHVSQTKES	ELAAIRNQKK	AELEAAKKAR
850	860	870	880	890	900	910
IEEVRKRRYE	EAIARRKEEI	ANAERQKRAQ	ELAEATRKQR	EIEEAAAKKS	TPYSFRAGNR	EPPSTPSTLP
920	930	940	950	960	970	980
KATVSPDKAK	LDMIAQKQRE	MEEAIEQRLA	GRTAGGSSPA	TPATPATPAT	PTPSSGPKKM	TMAEKLRAKR

LAKGGR

eIF3aN1-His6 (MW = 108kDa ; $\epsilon = 0,648\text{cm}^2/\text{mg}$; pI = 6,33) pET15bIpa-3aN1

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MLEAPPPFRP	ENAIKRADEL	ISVGEKQAAL	QSLHDFITAR	RIRWATPSTV
80	90	100	110	120	130	140
EPVVFKFLEI	GVELKKGKLL	KDGLHQYKLL	IQGSTEGLVS	VGAVARKFID	LVESKIASAQ	TRADELQKQE
150	160	170	180	190	200	210
IDDDLEGGVT	PENLLISVYE	SDQSVAGFND	EAITSWLRFT	WESYRAVLDL	LRNNALLEIT	YSGVVKTMH
220	230	240	250	260	270	280
FCLKYQRKNE	FKRLAEMLRQ	HLDAANYQQS	KSGNNLVDLS	DADTLQRYLD	QRFQQVDVSV	KLELWHEAYR
290	300	310	320	330	340	350
SIEDVFHLMK	ISKRAPKPST	LANYYENLVK	VFFVSGDPLL	HTTAWKKFYK	LYSTNPRATE	EEFKTYSSTI
360	370	380	390	400	410	420
FLSAISTQLD	EIPSIGYDPH	LRMYRLLNLD	AKPTRKEMLQ	SIIEDESIYG	KVDEELKELY	DIIEVNFVDV
430	440	450	460	470	480	490
TVKQQLLENLL	VKLSSKTYFS	QYIAPLRDVI	MRRVFVAASQ	KFTTVSQSEL	YKLATLPAPL	DLSAWDIEKS
500	510	520	530	540	550	560
LLQAAVEDYV	SITIDHESAK	VTFADKPPFDI	FASTASKEVS	EEENTEPEVQ	EEKEETDEAL	GPQETEDGEE
570	580	590	600	610	620	630
KEEESDPVII	RNSYIHNKLL	ELSNVLHDVD	SFNNASYMEK	VRIARETLIK	KNKDDLEKIS	KIVDERVKRS
640	650	660	670	680	690	700
QEQQKQKHEH	AALHAEQDAE	VRQQRILEEK	AAIEAKLEEE	AHRRLIEKKK	REFEAIKERE	ITKMITEVNA
710	720	730	740	750	760	770
KGHVYIDPNE	AKSLDLDTIK	QVIIAEVSKN	KSELESMEY	AMKKLDHTE	ALRKVELPLL	QKEVDKLQET
780	790	800	810	820	830	840
DTANYEAMKK	KIVDAAKAEY	EARMADRKNL	VMVYDDYLKF	KEHVSQTKES	ELAAIRNQKK	AELEAAKKAR
850	860	870	880	890	900	910
IEEVRKRRYE	EAIARRKEEI	ANAERQKRAQ	ELAEATRKQR	EIEEAAAKKS	TPYSFRAGNR	EPPSTPSTLP
920	930	940				
KATVSPDKAK	LDMIAQKQRE	MEEAIEQRLA				

3aN2-His6**(MW = 102kDa ; ϵ = 0,674cm²/mg ; pI = 6,14) pET15bpa-3aN2**

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MLEAPPPFRP	ENAIKRADEL	ISVGEKQAAL	QSLHDFITAR	RIRWATPSTV
80	90	100	110	120	130	140
EPVVFKELEI	GVELKKGKLL	KDGLHQYKKL	IQGSTEGLVS	VGAVARKFID	LVESKIASEQ	TRADELQKQE
150	160	170	180	190	200	210
IDDDLEGGVT	PENLLISVYE	SDQSVAGFND	EAITSWLRFT	WESYRAVLDL	LRNNALLEIT	YSGVVKKTMH
220	230	240	250	260	270	280
FCLKYQRKNE	FKRLAEMLRQ	HLDAANYQQS	KSGNNLVDLS	DADTLQRYLD	QRFQQVDVSV	KLELWHEAYR
290	300	310	320	330	340	350
SIEDVFLMK	ISKRAPKPST	LANYYENLVK	VFFVSGDLL	HTTAWKKFYK	LYSTNPRATE	EEFKTYSSTI
360	370	380	390	400	410	420
FLSAISTQLD	EIPSIGYDPH	LRMYRLNLD	AKPTRKEMLQ	SIIEDESIYG	KVDEELKELY	DIIEVNFVDV
430	440	450	460	470	480	490
TVKQQLLENLL	VKLSSKTYFS	QYIAPLRDVI	MRRVFVAASQ	KFTTVSQSEL	YKLATLPAPL	DLSAWDIEKS
500	510	520	530	540	550	560
LLQAAVEDYV	SITIDHESAK	VTFKDPFDI	FASTASKEVS	EEENTEPEVQ	EEKEETDEAL	GPQETEDGEE
570	580	590	600	610	620	630
KEEESDPVII	RNSYIHNKLL	ELSNVLHDVD	SFNNASYMEK	VRIARETLIK	KNKDDLEKIS	KIVDERVKRS
640	650	660	670	680	690	700
QEQKQKHMEH	AALHAEQDAE	VRQQRILEEK	AAIEAKLEEE	AHRRLIEKKK	REFEAIKERE	ITKMITEVNA
710	720	730	740	750	760	770
KGHVYIDPNE	AKSLDLDTIK	QVIAAEVSKN	KSELESMEY	AMKKLDHTER	ALRKVELPLL	QKEVDKLQET
780	790	800	810	820	830	840
DTANYEAMKK	KIVDAAKAEY	EARMADRKNL	VMVYDDYLKF	KEHVSGTKES	ELAAIRNQQK	AELEAAKKAR
850	860	870	880			
IEEVRKRRYE	EAIARRKEEI	ANAERQKRAQ	ELAEATRKQR	EIEEAAA		

eIF3aN2-His6**(MW = 102kDa ; ϵ = 0,677cm²/mg ; pI = 5,98) pCDF-3aN2-***

10	20	30	40	50	60	70
MGSSHHHHHH	SQDPMLEAPP	PFRPENAIKR	ADELISVGEK	QAALQSLHDF	ITARRIRWAT	PSTVEPVVFK
80	90	100	110	120	130	140
FLEIGVELKK	GKLLKDGLHQ	YKKLIQGSTE	GLVSVGAVAR	KFIDLVESKI	ASEQTRADEL	QKQEIDDDLE
150	160	170	180	190	200	210
GGVTPENLLI	SVYESDQSV	GFNDEAITSW	LRFTWESYRA	VLDLLRNNAL	LEITYSGVVK	KTMHFCLKYQ
220	230	240	250	260	270	280
RKNEFKRLAE	MLRQHDAAN	YQQSKSGNNL	VDSLADATLQ	RYLDQRFOQV	DVSVKLELWH	EAYRSIEDVF
290	300	310	320	330	340	350
HLMKISKRAP	KPSTLANYYE	NLVKVVFFVSG	DPLLHTTAWK	KFYKLYSTNP	RATEEEFKTY	SSTIFLSAIS
360	370	380	390	400	410	420
TQLDEIPSIG	YDPHLRMYRL	LNLDAKPTRK	EMLQSIIEDE	SIYGKVDEEL	KELYDIEVN	FDVDTVKQQL
430	440	450	460	470	480	490
ENLLVKLSSK	TYFSQYIAPL	RDVIMRRVVF	AASQKFTTVS	QSELYKLATL	PAPLDLSAWD	IEKSLQAAV
500	510	520	530	540	550	560
EDYVSITIDH	ESAKVTFKAD	PFDIFASTAS	KEVSEEENTE	PEVQEEKEET	DEALGPQETE	DGEEKEEESD
570	580	590	600	610	620	630
PVIIRNSYIH	NKLELSNVL	HDVDSFNNAS	YMEKVRIARE	TLIKKNKDDL	EKISKIVDER	VKRSQEQKQK
640	650	660	670	680	690	700
HMEHAALHAE	QDAEVRQQRI	LEEKAAIEAK	LEEEAHRRLI	EKKKREFEAI	KEREITKMIT	EVNAKGHVYI

710	720	730	740	750	760	770
DPNEAKSLDL	DTIKQVIAE	VSKNKSELES	RMEYAMKKLD	HTERALRKVE	LPLLQKEVDK	LQETDTANYE
780	790	800	810	820	830	840
AMKKKIVDAA	KAEYEARMAD	RKNLVMVYDD	YLKFKEHVSG	TKESELAAIR	NQKKAEELEAA	KKARIEEVRK
850	860	870	880	890	900	910
RRYEEAIARR	KEEIANAERQ	KRAQELAEAT	RKQREIEEAA	A		

eIF3a-His6 (MW = 112kDa ; $\epsilon = 0,625\text{cm}^2/\text{mg}$; pI = 6,66) pCDF-3a-*

10	20	30	40	50	60	70
MGSSHHHHHH	SQDPMLEAPP	PFRPENAIKR	ADELISVGEK	QAALQSLHDF	ITARRIRWAT	PSTVEPVVFK
80	90	100	110	120	130	140
FLEIGVELKK	GKLLKDGLHQ	YKKLIQGSTE	GLVSVGAVAR	KFIDLVESKI	ASEQTRADEL	QKQEIDDDLE
150	160	170	180	190	200	210
GGVTPENLLI	SVYESDQSV	GFNDEAITSW	LRFTWESYRA	VLDLLRNNAL	LEITYSGVVK	KTMHFCLKYQ
220	230	240	250	260	270	280
RKNEFKRLAE	MLRQHLLDAAN	YQQSKSGNNL	VDLSDADTLQ	RYLDQRFQQV	DVSVKLELWH	EAYRSIEDVF
290	300	310	320	330	340	350
HLMKISKRAP	KPSTLANYYE	NLVKVFFVSG	DPLLHTTAWK	KFYKLYSTNP	RATEEEFKTY	SSTIFLSAIS
360	370	380	390	400	410	420
TQLDEIPSIG	YDPHLRMYRL	LNLDAKPTRK	EMLQSIIEDE	SIYGKVDEEL	KELYDIEEVN	FDVDTVKQQL
430	440	450	460	470	480	490
ENLLVKLSSK	TYFSQYIAPL	RDVIMRRVVF	AASQKFTTVS	QSELYKLATL	PAPLDSAWD	IEKSLQAAV
500	510	520	530	540	550	560
EDYVSITIDH	ESAKVTFKAD	PFDIFASTAS	KEVSEEENTE	PEVQEEKEET	DEALGPQETE	DGEEKEEESD
570	580	590	600	610	620	630
PVIIRNSYIH	NKLELSNVL	HDVDSFNNAS	YMEKVRIARE	TLIKKNKDDL	EKISKIVDER	VKRSQEQKQK
640	650	660	670	680	690	700
HMEHAALHAE	QDAEVRQQRI	LEEKAAIEAK	LEEEAHRRLI	EKKKREFEAI	KEREITKMIT	EVNAKGHVYI
710	720	730	740	750	760	770
DPNEAKSLDL	DTIKQVIAE	VSKNKSELES	RMEYAMKKLD	HTERALRKVE	LPLLQKEVDK	LQETDTANYE
780	790	800	810	820	830	840
AMKKKIVDAA	KAEYEARMAD	RKNLVMVYDD	YLKFKEHVSG	TKESELAAIR	NQKKAEELEAA	KKARIEEVRK
850	860	870	880	890	900	910
RRYEEAIARR	KEEIANAERQ	KRAQELAEAT	RKQREIEEAA	AKKSTPYSFR	AGNREPPSTP	STLPKATVSP
920	930	940	950	960	970	980
DKAKLDMIAQ	KQREMEEAIE	QRLAGRTAGG	SSPATPATPA	TPATPTPSSG	PKKMTMAEKL	RAKRLAKGGR

eIF3aN1-His6 (MW = 108kDa ; $\epsilon = 0,651\text{cm}^2/\text{mg}$; pI = 6,16) pCDF-3aN1-*

10	20	30	40	50	60	70
MGSSHHHHHH	SQDPMLEAPP	PFRPENAIKR	ADELISVGEK	QAALQSLHDF	ITARRIRWAT	PSTVEPVVFK
80	90	100	110	120	130	140
FLEIGVELKK	GKLLKDGLHQ	YKKLIQGSTE	GLVSVGAVAR	KFIDLVESKI	ASEQTRADEL	QKQEIDDDLE
150	160	170	180	190	200	210
GGVTPENLLI	SVYESDQSV	GFNDEAITSW	LRFTWESYRA	VLDLLRNNAL	LEITYSGVVK	KTMHFCLKYQ
220	230	240	250	260	270	280
RKNEFKRLAE	MLRQHLLDAAN	YQQSKSGNNL	VDLSDADTLQ	RYLDQRFQQV	DVSVKLELWH	EAYRSIEDVF
290	300	310	320	330	340	350
HLMKISKRAP	KPSTLANYYE	NLVKVFFVSG	DPLLHTTAWK	KFYKLYSTNP	RATEEEFKTY	SSTIFLSAIS

360	370	380	390	400	410	420
TQLDEIPSIG	YDPHLRMYRL	LNLDAKPTRK	EMLQSIIEDE	SIYGKVDEEL	KELYDIEEVN	FDVDTVKQQL
430	440	450	460	470	480	490
ENLLVKLSSK	TYFSQYIAPL	RDVIMRRVFN	AASQKFTTVS	QSSELYKLATL	PAPLDLSAWD	IEKSLQAAV
500	510	520	530	540	550	560
EDYVSITIDH	ESAKVTFAKD	PFDIFASTAS	KEVSEEENTE	PEVQEEKEET	DEALGPQETE	DGEEKEEESD
570	580	590	600	610	620	630
PVIIRNSYIH	NKLELSNVL	HDVDSFNNAS	YMEKVRIARE	TLIKKNKDDL	EKISKIVDER	VKRSQEQKQK
640	650	660	670	680	690	700
HMEHAALHAE	QDAEVRQQRI	LEEKAAIEAK	LEEEAHRRLI	EKKKREFEAI	KEREITKMIT	EVNAKGHVYI
710	720	730	740	750	760	770
DPNEAKSLDL	DTIKQVIAE	VSKNKSELES	RMEYAMKKLD	HTERALRKVE	LPLLQKEVDK	LQETDTANYE
780	790	800	810	820	830	840
AMKKKIVDAA	KAEYEARMAD	RKNLVMVYDD	YLKFKHEVSG	TKESELAAIR	NQKKAEELEA	KKARIEEVRK
850	860	870	880	890	900	910
RRYEEAIARR	KEEIANAERQ	KRAQELAEAT	RKQREIEEAA	AKKSTPYSFR	AGNREPPSTP	STLPKATVSP
920	930					
DKAKLDMIAQ	KQREMEEAIE	QRLA				

Séquences d'eIF3b

eIF3b native (MW = 88kDa ; ϵ = 1,382cm²/mg ; pI = 6,21)

10	20	30	40	50	60	70
MKNFLPRTLK	NIYELYFNNI	SVHSIVSRNT	QLKRSKIIQM	TTETFEDIKL	EDIPVDDIDF	SDLEEYKVT
80	90	100	110	120	130	140
EEFNFDQYIV	VNGAPVIPS	KVPVLKKALT	SLFSKAGKVV	NMEFPIDEAT	GKTKGFLFVE	CGSMNDAKKI
150	160	170	180	190	200	210
IKSFH GKRLD	LKHRLFLYTM	KDVERYNSDD	FDTEFREPDM	PTFVPSSSLK	SWLMDDKVRD	QFVLQDDVKT
220	230	240	250	260	270	280
SVFWNSMFNE	EDSLVESREN	WSTNYVRFSP	KGTYLFSYHQ	QGVTAWGGPN	FDRLRRFYHP	DVRNSSVSPN
290	300	310	320	330	340	350
EKYLVTTFSTE	PIIVEEDNEF	SPFTKKNEGH	QLCIWDIASG	LLMATFPVIK	SPYLKWPLVR	WSYNDKYCAR
360	370	380	390	400	410	420
MVGDSLIVHD	ATKNFMPLA	KALKPSGIRD	FSFAPEGVKL	QPFRRNGDEPS	VLLAYWTPET	NNSACTATIA
430	440	450	460	470	480	490
EVPRGRVLKT	VNLVQVSNVT	LHWQNQAEFL	CFNVERHTKS	GKTQFSNLQI	CRLTERDIPV	EKVELKDSVF
500	510	520	530	540	550	560
EFGWEPHG NR	FVTISVHEVA	DMNYAIPANT	IRFYAPETKE	KTDVIKRWSL	VKEIPKTFAN	TVSWSPAGRF
570	580	590	600	610	620	630
VVVGALVGPN	MRRSDLQFYD	MDYPGEKNIN	DNNDVSASLK	DVAHPTYSAA	TNITWDPSGR	YVTAWSSSLK
640	650	660	670	680	690	700
HKVEHGYKIF	NIAGNLVKED	IIAGFKNFAW	RPRPASILSN	AERKKVRKNL	REWSAQFEEQ	DAMEADTAMR
710	720	730	740	750	760	
DLILHQRELL	KQWTEYREKI	GQEMEKS MNF	KIFDVQPEDA	SDDFTTIEEI	VEEVLEETKE	KVE

eIF3b-His6 (MW = 90kDa ; ϵ = 1,348cm²/mg ; pI = 6,68) pET15b1pa-eIF3b

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPR RGSH	MKNFLPRTLK	NIYELYFNNI	SVHSIVSRNT	QLKRSKIIQM	TTETFEDIKL
80	90	100	110	120	130	140
EDIPVDDIDF	SDLEEYKVT	EEFNFDQYIV	VNGAPVIPS	KVPVLKKALT	SLFSKAGKVV	NMEFPIDEAT
150	160	170	180	190	200	210
GKTKGFLFVE	CGSMNDAKKI	IKSFH GKRLD	LKHRLFLYTM	KDVERYNSDD	FDTEFREPDM	PTFVPSSSLK
220	230	240	250	260	270	280
SWLMDDKVRD	QFVLQDDVKT	SVFWNSMFNE	EDSLVESREN	WSTNYVRFSP	KGTYLFSYHQ	QGVTAWGGPN
290	300	310	320	330	340	350
FDRLRRFYHP	DVRNSSVSPN	EKYLVTTFSTE	PIIVEEDNEF	SPFTKKNEGH	QLCIWDIASG	LLMATFPVIK
360	370	380	390	400	410	420
SPYLKWPLVR	WSYNDKYCAR	MVGDSLIVHD	ATKNFMPLA	KALKPSGIRD	FSFAPEGVKL	QPFRRNGDEPS
430	440	450	460	470	480	490
VLLAYWTPET	NNSACTATIA	EVPRGRVLKT	VNLVQVSNVT	LHWQNQAEFL	CFNVERHTKS	GKTQFSNLQI
500	510	520	530	540	550	560
CRLTERDIPV	EKVELKDSVF	EFGWEPHG NR	FVTISVHEVA	DMNYAIPANT	IRFYAPETKE	KTDVIKRWSL

Annexe 2 : Séquence de la sous-unité eIF3b produite à partir des différentes constructions indiquées en italique. La masse molaire (MW), le coefficient d'extinction molaire(ϵ) et le point isolélectrique (pI) ont été calculés à partir des séquences dans DNAid ([Dardel and Bensoussan, 1988](#)). Les séquences apparaissant surlignées en gris sont excisées par l'action de la thrombine.

570	580	590	600	610	620	630
VKEIPKTFAN	TVSWSPAGRF	VVVGALVGPN	MRRSDLQFYD	MDYPGEKNIN	DNNDVSASLK	DVAHPTYSAA
640	650	660	670	680	690	700
TNITWDPSGR	YVTAWSSSLK	HKVEHGYKIF	NIAGNLVKED	IIAGFKNFAW	RPRPASILSN	AERKKVRKNL
710	720	730	740	750	760	770
REWSAQFEEQ	DAMEADTAMR	DLILHQRELL	KQWTEYREKI	GQEMEKS MNF	KIFDVQPEDA	SDDFTTIEEI
780						
VEEVLEETKE	KVE					

eIF3b-His6 (MW = 89kDa ; ϵ = 1,365cm²/mg ; pI = 6,50) pET32b-eIF3b

10	20	30	40	50	60	70
MKNFLPRTLK	NIYELYFN NI	SVHSIVSRNT	QLKRSKIIQM	TTETTFEDIKL	EDIPVDDIDF	SDLEEYKVT
80	90	100	110	120	130	140
EEFNFDQYIV	VNGAPVIPSA	KVPVLKKALT	SLFSKAGKVV	NMEFPIDEAT	GKTKGFLFVE	CGSMNDAKKI
150	160	170	180	190	200	210
IKSFH GKRLD	LKHRLFLYTM	KDVERYNSDD	FDTEFREPD M	PTFVPSSSLK	SWLMDDKVRD	QFVLQDDVKT
220	230	240	250	260	270	280
SVFWNSMFNE	EDSLVESREN	WSTNYVRFSP	KGTYLFSYHQ	QGV TAWGGPN	FDRLRRFYHP	DVRNSSVSPN
290	300	310	320	330	340	350
EKYLVTTFSTE	PIIVEEDNEF	SPFTKKNEGH	QLCIWDIASG	LLMATFPVIK	SPYLKWPLVR	WSYNDKYCAR
360	370	380	390	400	410	420
MVGDSLIVHD	ATKNFMPL EA	KALKPSGIRD	FSFAPEGVKL	QPFRRNGDEPS	VLLAYWTPET	NNSACTATIA
430	440	450	460	470	480	490
EVPRGRVLKT	VNLVQVSNVT	LHWQQA EFL	CFNVERHTKS	GKTQFSNLQI	CRLTERDIPV	EKVELKDSVF
500	510	520	530	540	550	560
EFGWEPHG NR	FVTISVHEVA	DMNYAIPANT	IRFYAPETKE	KTDVIKRWSL	VKEIPKTFAN	TVSWSPAGRF
570	580	590	600	610	620	630
VVVGALVGPN	MRRSDLQFYD	MDYPGEKNIN	DNNDVSASLK	DVAHPTYSAA	TNITWDPSGR	YVTAWSSSLK
640	650	660	670	680	690	700
HKVEHGYKIF	NIAGNLVKED	IIAGFKNFAW	RPRPASILSN	AERKKVRKNL	REWSAQFEEQ	DAMEADTAMR
710	720	730	740	750	760	770
DLILHQRELL	KQWTEYREKI	GQEMEKS MNF	KIFDVQPEDA	SDDFTTIEEI	VEEVLEETKE	KVELEHHHHH

H

eIF3bC1-His6 (MW = 67kDa ; ϵ = 1,632cm²/mg ; pI = 6,47) pET32b-3bC1

10	20	30	40	50	60	70
MDDKVRDQFV	LQDDVKTSVF	WNSMFNEEDS	LVESRENWST	NYVRFSPKGT	YLFSYHQQGV	TAWGGPNFDR
80	90	100	110	120	130	140
LRRFYHPDVR	NSSVSPNEKY	LVTTFSTEP II	VEEDNEFSPF	TKKNEGHQLC	IWDIASGLLM	ATFPVIKSPY
150	160	170	180	190	200	210
LKWPLVRWSY	NDKYCARMVG	DSLIVHDATK	NFMPL EAKAL	KPSGIRDFSF	APEGVKLQPF	RNGDEPSVLL
220	230	240	250	260	270	280
AYWTPETNNS	ACTATIAEVP	RGRVLKTVNL	VQVSNVTLHW	QNQA EFLCFN	VERHTKSGKT	QFSNLQICRL
290	300	310	320	330	340	350
TERDIPVEKV	ELKDSVF EFG	WEPHG NR FVT	ISVHEVADMN	YAIPANTIRF	YAPETKEKTD	VIKRWSLVKE
360	370	380	390	400	410	420
IPKTFANTVS	WSPAGRFVVV	GALVGPNMRR	SDLQFYDMDY	PGEKNINDNN	DVSASLKDVA	HPTYS AATNI

430	440	450	460	470	480	490
TWDPSSGRYVT	AWSSSLKHKV	EHGYKIFNIA	GNLVKEDIIA	GFKNFAWRPR	PASILSNAER	KKVRKNLREW
500	510	520	530	540	550	560
SAQFEEQDAM	EADTAMRDLI	LHQRELLKQW	TEYREKIGQE	MEKSMNFKIF	DVQPEDASDD	FTTIEEIVVEE
570	580					
VLEETKEKVE	LEHHHHHHH_					

eIF3bC1throm (MW = 67kDa ; ϵ = 1,617cm²/mg ; pI = 6,56) pET32b-3bC1thromb pCDF-3bC1-*

10	20	30	40	50	60	70
MDDKVRDQFV	LQDDVKTSVF	WNSMFNEEDS	LVESRENWST	NYVRFSPKGT	YLFSYHQQGV	TAWGGPNFDR
80	90	100	110	120	130	140
LRRFYHPDVR	NSSVSPNEKY	LVTTFSTPEII	VEEDNEFSPF	TKKNEGHQLC	IWDIASGLLM	ATFPVIKSPY
150	160	170	180	190	200	210
LKWPLVRWSY	NDKYCARMVG	DSLIVHDATAK	NFMPLAKAL	KPSGIRDFSF	APEGVKLQPF	RNGDEPSVLL
220	230	240	250	260	270	280
AYWTPETNNS	ACTATIAEVP	RGRVLKTVNL	VQVSNVTLHW	QNQAFLCFN	VERHTKSGKT	QFSNLQICRL
290	300	310	320	330	340	350
TERDIPVEKV	ELKDSVFEEG	WEPHGNERFVT	ISVHEVADMN	YAI PANTIRF	YAPETKEKTD	VIKRWSLVKE
360	370	380	390	400	410	420
IPKTFANTVS	WSPAGRFFVV	GALVGPNMRR	SDLQFYDMDY	PGEKNINDNN	DVSASLKDVA	HPTYSAATNI
430	440	450	460	470	480	490
TWDPSSGRYVT	AWSSSLKHKV	EHGYKIFNIA	GNLVKEDIIA	GFKNFAWRPR	PASILSNAER	KKVRKNLREW
500	510	520	530	540	550	560
SAQFEEQDAM	EADTAMRDLI	LHQRELLKQW	TEYREKIGQE	MEKSMNFKIF	DVQPEDASDD	FTTIEEIVVEE
570	580					
VLEETKEKVE	LLVPRGSEHH	HHHH				

eIF3bC2-His6 (MW = 21kDa ; ϵ = 1,529cm²/mg ; pI = 5,93) pET32b-3bC2

10	20	30	40	50	60	70
MDNNDVSASL	KDVAHPTYSA	ATNITWDPSG	RYVTAWSSSL	KHKVEHGYKI	FNIAGNLVKE	DIIAGFKNFA
80	90	100	110	120	130	140
WRPRPASILS	NAERKKVRKN	LREWSAQFEE	QDAMEADTAM	RDLILHQREL	LKQWTEYREK	IGQEMEKS MN
150	160	170	180			
FKIFDVQPED	ASDDFTTIEE	IVEEVLEETK	EKVELEHHHH	HH		

Séquences d'eIF3c

eIF3c native (MW = 93kDa ; $\epsilon = 0,774\text{cm}^2/\text{mg}$; pI = 5,09)

10	20	30	40	50	60	70
MSRFFSSNYE	YDVASSSSEE	DLLSSSEEDL	LSSSSSESEL	DQESDDSFN	ESESESEADV	DSDDSDAKPY
80	90	100	110	120	130	140
GPDWFKSEF	RKQGGGSKNF	LKSSNYDSSD	EESDEEDGKK	VVKSACEKLL	DEMQDVYNKI	SQAENSDDWL
150	160	170	180	190	200	210
TISNEFDLIS	RLLVRAQQQN	WGTPNIFIKV	VAQVEDAVNN	TQQADLKNKA	VARAYNTTKQ	RVKKVSRNE
220	230	240	250	260	270	280
DSMAKFRNDP	ESFDKEPTAD	LDISANGFTI	SSSQGNDQAV	QEDFFTRLQT	IIDSRGKKTV	NQQLSLISTLE
290	300	310	320	330	340	350
ELLTVAEKPY	EFIMAYLTLI	PSRFDASANL	SYQPIDQWKS	SFNDISKLLS	ILDQTIPTYQ	VNEFADPIDF
360	370	380	390	400	410	420
IEDEPKEDSD	GVKRILGSIF	SFVERLDDEF	MKSLLNIDPH	SSDYLRIRD	EQSIYNLILR	TQLYFEATLK
430	440	450	460	470	480	490
DEHDLERALT	RPFVKRLDHI	YYKSENLIKI	METAAWNIIIP	AQFKSKFTSK	DQLDSADYVD	NLIDGLSTIL
500	510	520	530	540	550	560
SKQNNIAVQK	RAILYNIYYT	ALNKDFQTAK	DMLLTSQVQT	NINQFDSSSLQ	ILFNRVVVQL	GLSAFKLCLI
570	580	590	600	610	620	630
EECHQILNDL	LSSSHLREIL	GQQLHRISL	NSSNNASADE	RARQCLPYHQ	HINLDLIDVV	FLTCSLLIEI
640	650	660	670	680	690	700
PRMTAFYSGI	KVKRIPYSPK	SIRRSLEHYD	KLSFQGPET	LRDYVLFAAK	SMQKGNWRDS	VKYLREIKSW
710	720	730	740	750	760	770
ALLPNMETVL	NSLTERVQVE	SLKTYFFSFK	RFYSSFSVAK	LAELFDLPEN	KVVEVLQSVI	AELEIPAKLN
780	790	800	810			
DEKTIFVVEK	GDEITKLEEA	MVKLNKEYKI	AKERLNPPSN	RR		

eIF3c-His6 (MW = 96kDa ; $\epsilon = 0,753\text{cm}^2/\text{mg}$; pI = 5,24) pET15blpa-eIF3c

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MPGLESRFFS	SNYEYDVASS	SSEEDLLSSS	EEDLLSSSSS	ESELDQESDD
80	90	100	110	120	130	140
SFFNESESES	EADVSDSD	AKPYGPDWFK	KSEFRKQGGG	SNKFLKSSNY	DSSDEESDEE	DGKKVVKSAK
150	160	170	180	190	200	210
EKLLDEMQDV	YNKISQAENS	DDWLTISNEF	DLISRLLVRA	QQQNWGTPNI	FIKVVAQVED	AVNNTQQADL
220	230	240	250	260	270	280
KNKAVARAYN	TTKQRVKKVS	RENEDSMAKF	RNDPESFDKE	PTADLDISAN	GFTISSSQGN	DQAVQEDFFT
290	300	310	320	330	340	350
RLQTIIDSRG	KKTVNQQLI	STLEELLTVA	EKPYEYFIMAY	LTLIPSRFDA	SANLSYQPID	QWKSSFNDIS
360	370	380	390	400	410	420
KLLSILDQTI	DTYQVNEFAD	PIDFIEDEPK	EDSDGVKRIL	GSIFSFVERL	DDEFMKSLLN	IDPHSSDYLI
430	440	450	460	470	480	490
RLRDEQSIYN	LILRTQLYFE	ATLKDEHDLE	RALTRPFVKR	LDHIYYKSEN	LIKIMETAAW	NIIPAQFKSK

Annexe 3 : Séquence de la sous-unité eIF3c produite à partir des différentes constructions indiquées en italique. La masse molaire (MW), le coefficient d'extinction molaire (ϵ) et le point isolélectrique (pI) ont été calculés à partir des séquences dans DNAid ([Dardel and Bensoussan, 1988](#)). Les séquences apparaissant surlignées en gris sont excisées par l'action de la thrombine.

500	510	520	530	540	550	560
FTSKDQLDSA	DYVDNLIDGL	STILSKQNNI	AVQKRAILYN	IYYTALNKDF	QTAKDMLLTS	QVQTININQFD
570	580	590	600	610	620	630
SSLQILFNRV	VVQLGLSAFK	LCLIEECHQI	LNDLLSSSHL	REILGQQLSH	RISLNSSNNA	SADERARQCL
640	650	660	670	680	690	700
PYHQHINLDL	IDVVFLTCSL	LIEIPRMTAF	YSGIKVKRIP	YSPKSIRRS	EHYDKLSFQG	PPETLRDYVL
710	720	730	740	750	760	770
FAAKSMQKGN	WRDSVKYLRE	IKSWALLPNM	ETVLNSLTER	VQVESLKYTF	FSFKRFYSSF	SVAKLAELFD
780	790	800	810	820	830	
LPENKVVEVL	QSVIAELEIP	AKLNDEKTIF	VVEKGDEITK	LEEAMVKLNK	EYKIAKERLN	PPSNRR

eIF3cC1-His6 (MW = 78kDa ; $\epsilon = 0,708\text{cm}^2/\text{mg}$; pI = 6.73) pET15blpa-eIF3cC1

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MQQNWGTPNI	FIKVVAQVED	AVNNTQQADL	KNKAVARAYN	TTKQRVKKVS
80	90	100	110	120	130	140
RENEDSMAKF	RNDPESFDKE	PTADLDISAN	GFTISSSQGN	DQAVQEDFFT	RLQTIIDSRG	KKTVNQQLSI
150	160	170	180	190	200	210
STLEELLTVA	EKPYEFIMAY	LTLIPSRFDA	SANLSYQPID	QWKSSFNDIS	KLLSILDQTI	DTYQVNEFAD
220	230	240	250	260	270	280
PIDFIEDEPK	EDSDGVKRIL	GSIFSFVERL	DDEFMKSLLN	IDPHSSDYLI	RLRDEQSIYN	LILRTQLYFE
290	300	310	320	330	340	350
ATLKDEHDLE	RALTRPFVKR	LDHIYYKSEN	LIKIMETAAW	NIIPAQFKSK	FTSKDQLDSA	DYVDNLIDGL
360	370	380	390	400	410	420
STILSKQNNI	AVQKRAILYN	IYYTALNKDF	QTAKDMLLTS	QVQTININQFD	SSLQILFNRV	VVQLGLSAFK
430	440	450	460	470	480	490
LCLIEECHQI	LNDLLSSSHL	REILGQQLSH	RISLNSSNNA	SADERARQCL	PYHQHINLDL	IDVVFLTCSL
500	510	520	530	540	550	560
LIEIPRMTAF	YSGIKVKRIP	YSPKSIRRS	EHYDKLSFQG	PPETLRDYVL	FAAKSMQKGN	WRDSVKYLRE
570	580	590	600	610	620	630
IKSWALLPNM	ETVLNSLTER	VQVESLKYTF	FSFKRFYSSF	SVAKLAELFD	LPENKVVEVL	QSVIAELEIP
640	650	660	670			
AKLNDEKTIF	VVEKGDEITK	LEEAMVKLNK	EYKIAKERLN	PPSNRR		

eIF3c (MW = 95kDa ; $\epsilon = 0,814\text{cm}^2/\text{mg}$; pI = 5,06) pCDF-*-3c

10	20	30	40	50	60	70
MADLNWISAG	HAIADVGTMP	GLESRFFSSN	YEYDVASSSS	EEDLLSSSEE	DLLSSSSSES	ELDQESDDSF
80	90	100	110	120	130	140
FNESESESEA	DVDSDDSDAK	PYGPWFKKS	EFRKQGGGSN	KFLKSSNYDS	SDEESDEEDG	KKVVKSAKEK
150	160	170	180	190	200	210
LLDEMQDVYN	KISQAENSDD	WLTISNEFDL	ISRLLVRAQQ	QNWGTPNIFI	KVVAQVEDAV	NNTQQADLKN
220	230	240	250	260	270	280
KAVARAYNTT	KQRVKKVSRE	NEDSMAKFRN	DPESFDKEPT	ADLDISANGF	TISSSQGNDQ	AVQEDFFTRL
290	300	310	320	330	340	350
QTIIDSRGKK	TVNQQLIST	LEELLTVAEK	PYEFIMAYLT	LIPSRFDASA	NLSYQPIDQW	KSSFNDISKL
360	370	380	390	400	410	420
LSILDQTIDT	YQVNEFADPI	DFIEDEPKED	SDGVKRILGS	IFSFVERLDD	EFMKSLLNID	PHSSDYLRIL
430	440	450	460	470	480	490
RDEQSIYNLI	LRTQLYFEAT	LKDEHDLERA	LTRPFVKRLD	HIYYKSENLI	KIMETAAWNI	IPAQFKSKFT

500	510	520	530	540	550	560
SKDQLDSADY	VDNLIDGLST	ILSKQNNIAV	QKRAILYNIY	YTALNKDFQT	AKDMLLTSQV	QTNINQFDSS
570	580	590	600	610	620	630
LQILFNRVVV	QLGLSAFKLC	LIEECHQILN	DLLSSSHLRE	ILGQQSLHRI	SLNSSNNASA	DERARQCLPY
640	650	660	670	680	690	700
HQHINLDLID	VVFLTCSLLI	EIPRMTAFYS	GIKVKRIPYS	PKSIRRSLEH	YDKLSFQGPP	ETLRDYVLFA
710	720	730	740	750	760	770
AKSMQKGNWR	DSVKYLREIK	SWALLPNMET	VLNSLTERVQ	VESLKTYFFS	FKRFYSSFSV	AKLAELFDLP
780	790	800	810	820	830	
ENKVVEVLQS	VIAELEIPAK	LNDEKTIFVV	EKGDEITKLE	EAMVKLNKEY	KIAKERLNPP	SNRR

eIF3cN1 (MW = 23kDa ; ϵ = 1,034cm²/mg ; pI = 4,56) pCDF-*-3cN1

10	20	30	40	50	60	70
SRFFSSNYEY	DVASSSSEED	LLSSSEEDLL	SSSSSESELD	QESDDSFNE	SESESEADVD	SDDSDAKPYG
80	90	100	110	120	130	140
PDWFKKSEFR	KQGGGSNKFL	KSSNYDSSDE	ESDEEDGKKV	VKSAKEKLLD	EMQDVYNKIS	QAENSDDWLT
150	160	170	180	190	200	
ISNEFDLISR	LLVRAQQQNW	GTPNIFIKVV	AQVEDAVNNT	QQADLKNKAV	ARAYNTTKQR	VKKV

eIF3cC1 (MW = 77kDa ; ϵ = 0,784cm²/mg ; pI = 5,94) pCDF-*-3cC1

10	20	30	40	50	60	70
MADLNWISAG	HAIADVGTQQ	NWGTPNIFIK	VVAQVEDAVN	NTQQADLKNK	AVARAYNTTK	QRVKKVSREN
80	90	100	110	120	130	140
EDSMAKFRND	PESFDKEPTA	DLDISANGFT	ISSSQGNDQA	VQEDFFTRLQ	TIIDSRGKKT	VNQQSLISTL
150	160	170	180	190	200	210
EELLTVAEKP	YEFIMAYLTL	IPSRFDASAN	LSYQPIDQWK	SSFNDISKLL	SILDQTIDTY	QVNEFADPID
220	230	240	250	260	270	280
FIEDEPKEDS	DGVKRILGSI	FSFVERLDDE	FMKSLLNIDP	HSSDYLIIRLR	DEQSIYNLIL	RTQLYFEATL
290	300	310	320	330	340	350
KDEHDLERAL	TRPFVKRLDH	IYYKSENLIK	IMETAAWNII	PAQFKSKFTS	KDQLDSADYV	DNLIDGLSTI
360	370	380	390	400	410	420
LSKQNNIAVQ	KRAILYNIYY	TALNKDFQTA	KDMLLTSQVQ	TNINQFDSSL	QILFNRVVVQ	LGLSAFKLCL
430	440	450	460	470	480	490
IEECHQILND	LLSSSHLREI	LGQQSLHRIS	LNSSNNASAD	ERARQCLPYH	QHINLDLIDV	VFLTCSLLEI
500	510	520	530	540	550	560
IPRMTAFYSG	IKVKRIPYSP	KSIRRSLEHY	DKLSFQGPPE	TLRDYVLFAA	KSMQKGNWRD	SVKYLREIKS
570	580	590	600	610	620	630
WALLPNMETV	LNLSLTERVQV	ESLKTYFFSF	KRFYSSFSVA	KLAELFDLPE	NKVVEVLQSV	IAELEIPAKL
640	650	660	670			
NDEKTIFVVE	KGDEITKLEE	AMVKLNKEYK	IAKERLNPPS	NRR		

Séquences d'eIF3g

eIF3g native (MW = 30kDa ; $\epsilon = 0,780\text{cm}^2/\text{mg}$; pI = 7,07)

10	20	30	40	50	60	70
MSEVAPEEII	ENADGSRsii	TYKIEDGVKY	KITQKVKEVK	VLEKVKHSVA	ERKNWHKYGS	EKGSPAGPSA
80	90	100	110	120	130	140
VTARLGEEVE	LRLSRNWKQA	EEERIQKEKA	SLTKTGLQCR	LCGNDHMTMN	CPFKTILSEL	SALEDPATNE
150	160	170	180	190	200	210
GGVEAASEEK	AGQVGGAGSI	PGQYVPPSRR	AGARDPSSDA	YRDSRERDDM	CTLKIMQVNE	NADENSLREE
220	230	240	250	260	270	
LLFPFAPIPR	VSVVRNKETG	KSRGLAFVTF	SSEEVAEQAL	RFLDGRGYMN	LILRVEWSKP	KVKE

eIF3g-His6 (MW = 33kDa ; $\epsilon = 0,719\text{cm}^2/\text{mg}$; pI = 7,58) pET15blpa-eIF3g

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLV PRGSH	MPGLESEVAP	EIIENADGS	RSIITYKIED	GVKYKITQKV	KEVKVLEKVVH
80	90	100	110	120	130	140
KSVAERKNWH	KYGSEKSPA	GPSAVTARLG	EEVELRLSRN	WKQAEERIQ	KEKASLTKTG	LQCRLCGNDH
150	160	170	180	190	200	210
MTMNCPFKTI	LSELSALEDP	ATNEGGVEAA	SEEKAGQVGG	AGSIPGQYVP	PSRRAGARDP	SSDAYRDSRE
220	230	240	250	260	270	280
RDDMCTLKIM	QVNENADENS	LREELLFPFA	PIPRVSVVRN	KETGKSRGLA	FVTFSSSEVA	EQALRFLDGR
290						
GYMNLILRVE	WSKPKVKE					

eIF3gC1-His6 (MW = 28kDa ; $\epsilon = 0,763\text{cm}^2/\text{mg}$; pI = 8,18) pET15blpa-eIF3gC1

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLV PRGSH	MEKVHKSVAE	RKNWHKYGSE	KGSPAGPSAV	TARLGEEVEL	RLSRNWKQAE
80	90	100	110	120	130	140
EERIQKEKAS	LTKTGLQCRL	CGNDHMTMNC	PFKTIILSLS	ALEDPATNEG	GVEAASEEKA	GQVGGAGSIP
150	160	170	180	190	200	210
GQYVPPSRRRA	GARDPSSDAY	RDSRERDDMC	TLKIMQVNEN	ADENSLREEL	LFPFAPIPRV	SVVRNKETGK
220	230	240	250			
SRGLAFVTF	SEEVAEQALR	FLDGRGYMNL	ILRVEWSKPK	VKE		

Annexe 4 : Séquence de la sous-unité eIF3g produite à partir des différentes constructions indiquées en italique. La masse molaire (MW), le coefficient d'extinction molaire (ϵ) et le point isolélectrique (pI) ont été calculés à partir des séquences dans DNAid ([Dardel and Bensoussan, 1988](#)). Les séquences apparaissant surlignées en gris sont excisées par l'action de la thrombine.

eIF3g (MW = 31kDa ; ϵ = 0,943cm²/mg ; pI = 6,65) pCOLA-*-3g

10	20	30	40	50	60	70
MADLNWSEVA	PEEIIENADG	SRSIIITYKIE	DGVKYKITQK	VKEVKVLEKV	HKSVAERKNW	HKYGSEKSP
80	90	100	110	120	130	140
AGPSAVTARL	GEEVELRLSR	NWKQAEERI	QKEKASLTKT	GLQCRLCGND	HMTMNCPFKT	ILSELSALED
150	160	170	180	190	200	210
PATNEGGVEA	ASEEKAGQVG	GAGSIPGQYV	PPSRRAGARD	PSSDAYRDSR	ERDDMCTLKI	MQVNENADEN
220	230	240	250	260	270	
SLREELLFPF	APIPRVSVVR	NKETGKSRGL	AFVTFSSEEV	AEQALRFLDG	RGYMNLI LRV	EWSKPKVKE

eIF3gC1 (MW = 26kDa ; ϵ = 0,827cm²/mg ; pI = 7,59) pCOLA-*-3gC1

10	20	30	40	50	60	70
MEKVHKSVAE	RKNWHKYGSE	KGSPAGPSAV	TARLGEEVEL	RLSRNWKQAE	EERIQKEKAS	LTKTGLQCRL
80	90	100	110	120	130	140
CGNDHMTMNC	PFKTI LSELS	ALEDPATNEG	GVEAASEEKA	GQVGGAGSIP	GQYVPPSRA	GARDPSSDAY
150	160	170	180	190	200	210
RDSRERDDMC	TLKIMQVNEN	ADENSLREEL	LFPFAPIPRV	SVVRNKETGK	SRGLAFVTFS	SEEVAEQALR
220	230					
FLDGRGYMNL	ILRVEWSKPK	VKE				

Séquences d'eIF3i

eIF3i native (MW = 39kDa ; ϵ = 1,207cm²/mg ; pI = 5,83) pET3alpa-eIF3i

10	20	30	40	50	60	70
MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG	HTGTIWSIDV	DCFTKYCVTG
80	90	100	110	120	130	140
SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNPG	SINIYEIERD	SATHELTKVS
150	160	170	180	190	200	210
EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD	SIDLHEKISIS	DMQFSPDLTY
220	230	240	250	260	270	280
FITSSRDNTS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIIILGGGQEA	KDVTTTSANE	GKFEARFYHK
290	300	310	320	330	340	350
IFEEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYDFDK	YDVEKAAEAK	EHMQEAN

eIF3i-His6 (MW = 41kDa ; ϵ = 1,143cm²/mg ; pI = 6,59) pET15blpa-eIF3i

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG
80	90	100	110	120	130	140
HTGTIWSIDV	DCFTKYCVTG	SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNPG
150	160	170	180	190	200	210
SINIYEIERD	SATHELTKVS	EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD
220	230	240	250	260	270	280
SIDLHEKISIS	DMQFSPDLTY	FITSSRDNTS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIIILGGGQEA
290	300	310	320	330	340	350
KDVTTTSANE	GKFEARFYHK	IFEEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYDFDK
360						
YDVEKAAEAK	EHMQEAN					

eIF3i-His6 (MW = 40kDa ; ϵ = 1,159cm²/mg ; pI = 6,34) pCOLA-3i-*

10	20	30	40	50	60	70
MGSSHHHHHH	SQDPMKAIKL	TGHERPLTQV	KYNKEGDLLE	SCSKDSSASV	WYSLNGERLG	TLDGHTGTIW
80	90	100	110	120	130	140
SIDVDCFTKY	CVTGSADYSI	KLWDVSNGQC	VATWKSPVPV	KRVEFSPCGN	YFLAILDNVM	KNPGSINIYE
150	160	170	180	190	200	210
IERDSATHEL	TKVSEEPIHK	IITHEGLDAA	TVAGWSTKGK	YIIAGHKDGK	ISKYDVSNNY	EYVDSIDLHE
220	230	240	250	260	270	280
KSISDMQFSP	DLTYFITSSR	DTNSFLVDVS	TLQVLKKYET	DCPLNTAVIT	PLKEFIIILGG	GQEAQDVTTT
290	300	310	320	330	340	350
SANEGKFEAR	FYHKIFEEEEI	GRVQGHFGPL	NTVAISPQGT	SYASGGEDGF	IRLHHFEKSY	DFDKYDVEKA
360						
AEAKEHMQEA	N					

Annexe 5 : Séquence de la sous-unité eIF3i produite à partir des différentes constructions indiquées en italique. La masse molaire (MW), le coefficient d'extinction molaire(E) et le point isolélectrique (pI) ont été calculés à partir des séquences dans DNAid ([Dardel and Bensoussan, 1988](#)). Les séquences apparaissant surlignées en gris sont excisées par l'action de la thrombine.

eIF3i-His6 (MW = 48kDa ; ϵ = 1,157cm²/mg ; pI = 6,44) pET32b-eIF3i

10	20	30	40	50	60	70
MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG	HTGTIWSIDV	DCFTKYCVTG
80	90	100	110	120	130	140
SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNP	SINIYEIERD	SATHELTKVS
150	160	170	180	190	200	210
EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD	SIDLHEKSIS	DMQFSPDLTY
220	230	240	250	260	270	280
FITSSRDNTS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIILGGGQEA	KDVTTTSANE	GKFEARFYHK
290	300	310	320	330	340	350
IFEEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYFDFK	YDVEKAAEAK	EHMQEANLVP
360						
RGSLEHHHHH	H					

eIF3iDEL1 (MW = 39kDa ; ϵ = 1,193cm²/mg ; pI = 6,57) pET32b-3iDEL1

10	20	30	40	50	60	70
MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG	HTGTIWSIDV	DCFTKYCVTG
80	90	100	110	120	130	140
SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNP	SINIYEIERD	SATHELTKVS
150	160	170	180	190	200	210
EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD	SIDLHEKSIS	DMQFSPDLTY
220	230	240	250	260	270	280
FITSSRDNTS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIILGGGQEA	KDVTTTSANE	GKFEARFYHK
290	300	310	320	330	340	350
IFEEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYFDFK	YDVEKALVPR	GSLEHHHHHH

eIF3iDEL2 (MW = 40kDa ; ϵ = 1,170cm²/mg ; pI = 6,53) pET32b-eIF3iDEL2

10	20	30	40	50	60	70
MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG	HTGTIWSIDV	DCFTKYCVTG
80	90	100	110	120	130	140
SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNP	SINIYEIERD	SATHELTKVS
150	160	170	180	190	200	210
EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD	SIDLHEKSIS	DMQFSPDLTY
220	230	240	250	260	270	280
FITSSRDNTS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIILGGGQEA	KDVTTTSANE	GKFEARFYHK
290	300	310	320	330	340	350
IFEEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYFDFK	YDVEKAAEAK	EHMLVPRGSL

EHHHHHHH**eIF3iSTOP1 (MW = 40kDa ; ϵ = 1,167cm²/mg ; pI = 6,72) pET15blpa-3iSTOP1**

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG
80	90	100	110	120	130	140
HTGTIWSIDV	DCFTKYCVTG	SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNP
150	160	170	180	190	200	210
SINIYEIERD	SATHELTKVS	EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD

220	230	240	250	260	270	280
SIDLHEKSI	DMQFSPDLTY	FITSSRDTNS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIIILGGGQEA
290	300	310	320	330	340	350
KDVTTSANE	GKFEARFYHK	IFEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYDFDK
360						
YDVEKAAEAK						

eIF3iSTOP2 (MW = 40kDa ; ϵ = 1,181cm²/mg ; pI = 6,72) pET15bpa-3iSTOP2

10	20	30	40	50	60	70
MGSSHHHHHH	SSGLVPRGSH	MKAIKLTGHE	RPLTQVKYNK	EGDLLFSCSK	DSSASVWYSL	NGERLGTLDG
80	90	100	110	120	130	140
HTGTIWSIDV	DCFTKYCVTG	SADYSIKLWD	VSNGQCVATW	KSPVPVKRVE	FSPCGNYFLA	ILDNVMKNPG
150	160	170	180	190	200	210
SINIYEIERD	SATHELTKVS	EEPIHKIITH	EGLDAATVAG	WSTKGKYIIA	GHKDGKISKY	DVSNNYEYVD
220	230	240	250	260	270	280
SIDLHEKSI	DMQFSPDLTY	FITSSRDTNS	FLVDVSTLQV	LKKYETDCPL	NTAVITPLKE	FIIILGGGQEA
290	300	310	320	330	340	350
KDVTTSANE	GKFEARFYHK	IFEEEIGRVQ	GHFGPLNTVA	ISPQGTSYAS	GGEDGFIRLH	HFEKSYDFDK

YDVEK

Caractéristiques des complexes purifiés

Complexe	MW (kDa)	ϵ (cm ² /mg)	pI
eIF3i-gC1	66	1,029	6,55
eIF3a-cC1	190	0,689	6,42
eIF3bC1-i	107	1,454	6,41
eIF3bC1-cC1	144	1,177	6,24
eIF3bC2-cC1	99	0,943	5,94
eIF3bC1-i-gC1	133	1,332	6,52
eIF3a-bC1-cC1	256	0,935	6,43
eIF3cC1-i-gC1	144	0,896	6,31
eIF3a-cC1-i-gC1	256	0,777	6,46
eIF3bC1-cC1-i-gC1	210	1,130	6,37
eIF3a-bC1-cC1-i-gC1	323	0,954	6,47

Annexe 5 : Caractéristiques des différents complexes purifiés. La masse molaire (MW), le coefficient d'extinction molaire (E) et le point isolélectrique (pI) ont été calculés à partir des séquences dans DNAid ([Dardel and Bensoussan, 1988](#)).

Références Bibliographiques

[Acker, M.G., Shin, B.S., Dever, T.E., and Lorsch, J.R. \(2006\). Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. *J Biol Chem* 281, 8469-8475.](#)

[Agrawal, R.K., Sharma, M.R., Kiel, M.C., Hirokawa, G., Booth, T.M., Spahn, C.M., Grassucci, R.A., Kaji, A., and Frank, J. \(2004\). Visualization of ribosome-recycling factor on the *Escherichia coli* 70S ribosome: functional implications. *Proc Natl Acad Sci U S A* 101, 8900-8905.](#)

[Akiyoshi, Y., Clayton, J., Phan, L., Yamamoto, M., Hinnebusch, A.G., Watanabe, Y., and Asano, K. \(2001\). Fission yeast homolog of murine Int-6 protein, encoded by mouse mammary tumor virus integration site, is associated with the conserved core subunits of eukaryotic translation initiation factor 3. *J Biol Chem* 276, 10056-10062.](#)

[Algire, M.A., and Lorsch, J.R. \(2006\). Where to begin? The mechanism of translation initiation codon selection in eukaryotes. *Curr Opin Chem Biol* 10, 480-486.](#)

[Algire, M.A., Maag, D., and Lorsch, J.R. \(2005\). Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell* 20, 251-262.](#)

[Allen, G.S., Zavialov, A., Gursky, R., Ehrenberg, M., and Frank, J. \(2005\). The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell* 121, 703-712.](#)

[Alone, P.V., and Dever, T.E. \(2006\). Direct binding of translation initiation factor eIF2gamma-G domain to its GTPase-activating and GDP-GTP exchange factors eIF5 and eIF2B epsilon. *J Biol Chem* 281, 12636-12644.](#)

[Altmann, M., Schmitz, N., Berset, C., and Trachsel, H. \(1997\). A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. *EMBO J* 16, 1114-1121.](#)

[Andreev, D.E., Terenin, I.M., Dunaevsky, Y.E., Dmitriev, S.E., and Shatsky, I.N. \(2006\). A leaderless mRNA can bind to mammalian 80S ribosomes and direct polypeptide synthesis in the absence of translation initiation factors. *Mol Cell Biol* 26, 3164-3169.](#)

[Antoun, A., Pavlov, M.Y., Andersson, K., Tenson, T., and Ehrenberg, M. \(2003\). The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO J* 22, 5593-5601.](#)

[Asano, K., Clayton, J., Shalev, A., and Hinnebusch, A.G. \(2000\). A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA\(Met\) is an important translation initiation intermediate *in vivo*. *Genes Dev* 14, 2534-2546.](#)

[Asano, K., Krishnamoorthy, T., Phan, L., Pavitt, G.D., and Hinnebusch, A.G. \(1999\). Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *EMBO J* 18, 1673-1688.](#)

[Asano, K., Phan, L., Anderson, J., and Hinnebusch, A.G. \(1998\). Complex formation by all five homologues of mammalian translation initiation factor 3 subunits from yeast *Saccharomyces cerevisiae*. *J Biol Chem* 273, 18573-18585.](#)

[Asano, K., Phan, L., Valasek, L., Schoenfeld, L.W., Shalev, A., Clayton, J., Nielsen, K., Donahue, T.F., and Hinnebusch, A.G. \(2001a\). A multifactor complex of eIF1, eIF2, eIF3, eIF5, and tRNA\(i\)Met promotes initiation complex assembly and couples GTP hydrolysis to AUG recognition. *Cold Spring Harb Symp Quant Biol* 66, 403-415.](#)

[Asano, K., Shalev, A., Phan, L., Nielsen, K., Clayton, J., Valasek, L., Donahue, T.F., and Hinnebusch, A.G. \(2001b\). Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO J* 20, 2326-2337.](#)

[Asano, K., Vornlocher, H.P., Richter-Cook, N.J., Merrick, W.C., Hinnebusch, A.G., and Hershey, J.W. \(1997\). Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits. Possible roles in RNA binding and macromolecular assembly. *J Biol Chem* 272, 27042-27052.](#)

[Bachmann, F., Banziger, R., and Burger, M.M. \(1997\). Cloning of a novel protein overexpressed in human mammary carcinoma. *Cancer Res* 57, 988-994.](#)

[Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. \(2000\). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905-920.](#)

[Bandyopadhyay, A., Matsumoto, T., and Maitra, U. \(2000\). Fission yeast Int6 is not essential for global translation initiation, but deletion of int6\(+\) causes hypersensitivity to caffeine and affects spore formation. Mol Biol Cell 11, 4005-4018.](#)

[Basu, U., Si, K., Warner, J.R., and Maitra, U. \(2001\). The *Saccharomyces cerevisiae* TIF6 gene encoding translation initiation factor 6 is required for 60S ribosomal subunit biogenesis. Mol Cell Biol 21, 1453-1462.](#)

[Battiste, J.L., Pestova, T.V., Hellen, C.U., and Wagner, G. \(2000\). The eIF1A solution structure reveals a large RNA-binding surface important for scanning function. Mol Cell 5, 109-119.](#)

[Baugh, J.M., and Pilipenko, E.V. \(2004\). 20S proteasome differentially alters translation of different mRNAs via the cleavage of eIF4F and eIF3. Mol Cell 16, 575-586.](#)

[Beckmann, R., Spahn, C.M., Eswar, N., Heliander, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. \(2001\). Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell 107, 361-372.](#)

[Bellsollell, L., Cho-Park, P.F., Poulin, F., Sonenberg, N., and Burley, S.K. \(2006\). Two structurally atypical HEAT domains in the C-terminal portion of human eIF4G support binding to eIF4A and Mnk1. Structure 14, 913-923.](#)

[Benne, R., Edman, J., Traut, R.R., and Hershey, J.W. \(1978\). Phosphorylation of eukaryotic protein synthesis initiation factors. Proc Natl Acad Sci U S A 75, 108-112.](#)

[Benne, R., and Hershey, J.W. \(1978\). The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. J Biol Chem 253, 3078-3087.](#)

[Berk, V., Zhang, W., Pai, R.D., and Cate, J.H. \(2006\). Structural basis for mRNA and tRNA positioning on the ribosome. Proc Natl Acad Sci U S A 103, 15830-15834.](#)

[Birnboim, H.C., and Doly, J. \(1979\). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7, 1513-1523.](#)

[Boesen, T., Mohammad, S.S., Pavitt, G.D., and Andersen, G.R. \(2004\). Structure of the catalytic fragment of translation initiation factor 2B and identification of a critically important catalytic residue. J Biol Chem 279, 10584-10592.](#)

[Boileau, G., Butler, P., Hershey, J.W., and Traut, R.R. \(1983\). Direct cross-links between initiation factors 1, 2, and 3 and ribosomal proteins promoted by 2-iminothiolane. Biochemistry 22, 3162-3170.](#)

[Boni, I.V., Isaeva, D.M., Musychenko, M.L., and Tzareva, N.V. \(1991\). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res 19, 155-162.](#)

[Bonneville, J.M., Sanfacon, H., Futterer, J., and Hohn, T. \(1989\). Posttranscriptional trans-activation in cauliflower mosaic virus. Cell 59, 1135-1143.](#)

[Borman, A.M., Michel, Y.M., and Kean, K.M. \(2000\). Biochemical characterisation of cap-poly\(A\) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5'-end. Nucleic Acids Res 28, 4068-4075.](#)

[Brevet, A., Chen, J., Leveque, F., Plateau, P., and Blanquet, S. \(1989\). In vivo synthesis of adenylylated bis\(5'-nucleosidyl\) tetraphosphates \(Ap4N\) by *Escherichia coli* aminoacyl-tRNA synthetases. Proc Natl Acad Sci U S A 86, 8275-8279.](#)

[Browning, K.S., Gallie, D.R., Hershey, J.W., Hinnebusch, A.G., Maitra, U., Merrick, W.C., and Norbury, C. \(2001\). Unified nomenclature for the subunits of eukaryotic initiation factor 3. Trends Biochem Sci 26, 284.](#)

[Bushell, M., Wood, W., Clemens, M.J., and Morley, S.J. \(2000\). Changes in integrity and association of eukaryotic protein synthesis initiation factors during apoptosis. Eur J Biochem 267, 1083-1091.](#)

[Buttitta, F., Martella, C., Barassi, F., Felicioni, L., Salvatore, S., Rosini, S., D'Antuono, T., Chella, A., Mucilli, F., Sacco, R., et al. \(2005\). Int6 expression can predict survival in early-stage non-small cell lung cancer patients. Clin Cancer Res 11, 3198-3204.](#)

[Caetano-Anolles, G. \(2002\). Tracing the evolution of RNA structure in ribosomes. *Nucleic Acids Res* 30, 2575-2587.](#)

[Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T., and Ramakrishnan, V. \(2000\). Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407, 340-348.](#)

[Carter, A.P., Clemons, W.M., Jr., Brodersen, D.E., Morgan-Warren, R.J., Hartsch, T., Wimberly, B.T., and Ramakrishnan, V. \(2001\). Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* 291, 498-501.](#)

[Caruthers, J.M., Johnson, E.R., and McKay, D.B. \(2000\). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc Natl Acad Sci U S A* 97, 13080-13085.](#)

[Castilho-Valavicius, B., Yoon, H., and Donahue, T.F. \(1990\). Genetic characterization of the *Saccharomyces cerevisiae* translational initiation suppressors *sui1*, *sui2* and *SUI3* and their effects on *HIS4* expression. *Genetics* 124, 483-495.](#)

[Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N., and Noller, H.F. \(1999\). X-ray crystal structures of 70S ribosome functional complexes. *Science* 285, 2095-2104.](#)

[Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L.A., Offenhauser, N., Marchisio, P.C., and Biffo, S. \(2003\). Release of eIF6 \(p27BBP\) from the 60S subunit allows 80S ribosome assembly. *Nature* 426, 579-584.](#)

[Chappell, S.A., Edelman, G.M., and Mauro, V.P. \(2000\). A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site \(IRES\) and when present in linked multiple copies greatly enhances IRES activity. *Proc Natl Acad Sci U S A* 97, 1536-1541.](#)

[Chaudhuri, J., Chakrabarti, A., and Maitra, U. \(1997\). 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, 30975-30983.](#)

[Chaudhuri, J., Chowdhury, D., and Maitra, U. \(1999\). Distinct functions of eukaryotic translation initiation factors eIF1A and eIF3 in the formation of the 40 S ribosomal preinitiation complex. *J Biol Chem* 274, 17975-17980.](#)

[Cheah, E., Carr, P.D., Suffolk, P.M., Vasudevan, S.G., Dixon, N.E., and Ollis, D.L. \(1994\). Structure of the *Escherichia coli* signal transducing protein PII. *Structure* 2, 981-990.](#)

[Chen, C.R., Li, Y.C., Chen, J., Hou, M.C., Papadaki, P., and Chang, E.C. \(1999\). Moe1, a conserved protein in *Schizosaccharomyces pombe*, interacts with a Ras effector, Scd1, to affect proper spindle formation. *Proc Natl Acad Sci U S A* 96, 517-522.](#)

[Chen, G., and Burger, M.M. \(1999\). p150 expression and its prognostic value in squamous-cell carcinoma of the esophagus. *Int J Cancer* 84, 95-100.](#)

[Chen, G., and Burger, M.M. \(2004\). p150 overexpression in gastric carcinoma: the association with p53, apoptosis and cell proliferation. *Int J Cancer* 112, 393-398.](#)

[Chen, R.H., Miettinen, P.J., Maruoka, E.M., Choy, L., and Derynck, R. \(1995\). A WD-domain protein that is associated with and phosphorylated by the type II TGF-beta receptor. *Nature* 377, 548-552.](#)

[Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. \(2003\). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31, 3497-3500.](#)

[Cheung, Y.N., Maag, D., Mitchell, S.F., Fekete, C.A., Algire, M.A., Takacs, J.E., Shirokikh, N., Pestova, T., Lorsch, J.R., and Hinnebusch, A.G. \(2007\). Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection *in vivo*. *Genes Dev* 21, 1217-1230.](#)

[Cho, S., and Hoffman, D.W. \(2002\). Structure of the beta subunit of translation initiation factor 2 from the archaeon *Methanococcus jannaschii*: a representative of the eIF2beta/eIF5 family of proteins. *Biochemistry* 41, 5730-5742.](#)

[Choi, S.K., Lee, J.H., Zoll, W.L., Merrick, W.C., and Dever, T.E. \(1998\). Promotion of met-tRNA^{iMet} binding to ribosomes by yIF2, a bacterial IF2 homolog in yeast. *Science* 280, 1757-1760.](#)

[Choi, S.K., Olsen, D.S., Roll-Mecak, A., Martung, A., Remo, K.L., Burley, S.K., Hinnebusch, A.G., and Dever, T.E. \(2000\). Physical and functional interaction between the eukaryotic orthologs of prokaryotic translation initiation factors IF1 and IF2. *Mol Cell Biol* 20, 7183-7191.](#)

[Choy, L., and Derynck, R. \(1998\). The type II transforming growth factor \(TGF\)-beta receptor-interacting protein TRIP-1 acts as a modulator of the TGF-beta response. *J Biol Chem* 273, 31455-31462.](#)

[Chudinova, E.M., Ivanov, P.A., and Nadezhdina, E.S. \(2004\). \[Large subunit of translation initiation factor--3 p170 contains potentially functional nuclear localization signals\]. *Mol Biol \(Mosk\)* 38, 684-691.](#)

[Clayton, R.A., White, O., Ketchum, K.A., and Venter, J.C. \(1997\). The first genome from the third domain of life. *Nature* 387, 459-462.](#)

[Clemens, M.J., Bushell, M., and Morley, S.J. \(1998\). Degradation of eukaryotic polypeptide chain initiation factor \(eIF\) 4G in response to induction of apoptosis in human lymphoma cell lines. *Oncogene* 17, 2921-2931.](#)

[Commans, S., Blanquet, S., and Plateau, P. \(1995\). A single substitution in the motif 1 of *Escherichia coli* lysyl-tRNA synthetase induces cooperativity toward amino acid binding. *Biochemistry* 34, 8180-8189.](#)

[Conte, M.R., Kelly, G., Babon, J., Sanfelice, D., Youell, J., Smerdon, S.J., and Proud, C.G. \(2006\). Structure of the eukaryotic initiation factor \(eIF\) 5 reveals a fold common to several translation factors. *Biochemistry* 45, 4550-4558.](#)

[Craig, A.W., Haghghat, A., Yu, A.T., and Sonenberg, N. \(1998\). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* 392, 520-523.](#)

[Crane, R., Craig, R., Murray, R., Dunand-Sauthier, I., Humphrey, T., and Norbury, C. \(2000\). A fission yeast homolog of Int-6, the mammalian oncoprotein and eIF3 subunit, induces drug resistance when overexpressed. *Mol Biol Cell* 11, 3993-4003.](#)

[Dahlquist, K.D., and Puglisi, J.D. \(2000\). Interaction of translation initiation factor IF1 with the *E. coli* ribosomal A site. *J Mol Biol* 299, 1-15.](#)

[Dallas, A., and Noller, H.F. \(2001\). Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Mol Cell* 8, 855-864.](#)

[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 Proteomics* 6, 1135-1146.](#)

[Danaie, P., Wittmer, B., Altmann, M., and Trachsel, H. \(1995\). Isolation of a protein complex containing translation initiation factor Prt1 from *Saccharomyces cerevisiae*. *J Biol Chem* 270, 4288-4292.](#)

[Dardel, F. \(1994\). MC-Fit: using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Comput Appl Biosci* 10, 273-275.](#)

[Dardel, F., and Bensoussan, P. \(1988\). DNAid: a Macintosh full screen editor featuring a built-in regular expression interpreter for the search of specific patterns in biological sequences using finite state automata. *Comput Appl Biosci* 4, 483-486.](#)

[Das, S., Ghosh, R., and Maitra, U. \(2001\). Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. *J Biol Chem* 276, 6720-6726.](#)

[Das, S., Maiti, T., Das, K., and Maitra, U. \(1997\). Specific interaction of eukaryotic translation initiation factor 5 \(eIF5\) with the beta-subunit of eIF2. *J Biol Chem* 272, 31712-31718.](#)

[Das, S., and Maitra, U. \(2000\). Mutational analysis of mammalian translation initiation factor 5 \(eIF5\): role of interaction between the beta subunit of eIF2 and eIF5 in eIF5 function *in vitro* and *in vivo*. *Mol Cell Biol* 20, 3942-3950.](#)

[Daughenbaugh, K.F., Fraser, C.S., Hershey, J.W., and Hardy, M.E. \(2003\). The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment. *EMBO J* 22, 2852-2859.](#)

[De Benedetti, A., and Graff, J.R. \(2004\). eIF-4E expression and its role in malignancies and metastases. *Oncogene* 23, 3189-3199.](#)

[Declercq, J.P., and Evrard, C. \(2001\). A twinned monoclinic crystal form of human peroxiredoxin 5 with eight molecules in the asymmetric unit. *Acta Crystallogr D Biol Crystallogr* 57, 1829-1835.](#)

[Dellas, A., Torhorst, J., Bachmann, F., Banziger, R., Schultheiss, E., and Burger, M.M. \(1998\). Expression of p150 in cervical neoplasia and its potential value in predicting survival. *Cancer* 83, 1376-1383.](#)

[Deo, R.C., Bonanno, J.B., Sonenberg, N., and Burley, S.K. \(1999\). Recognition of polyadenylate RNA by the poly\(A\)-binding protein. *Cell* 98, 835-845.](#)

[Desgres, J., Keith, G., Kuo, K.C., and Gehrke, C.W. \(1989\). Presence of phosphorylated O-ribosyl-adenosine in T-psi-stem of yeast methionine initiator tRNA. *Nucleic Acids Res* 17, 865-882.](#)

[Dever, T.E. \(2002\). Gene-specific regulation by general translation factors. *Cell* 108, 545-556.](#)

[Dever, T.E., and Hinnebusch, A.G. \(2005\). GCN2 whets the appetite for amino acids. *Mol Cell* 18, 141-142.](#)

[Dever, T.E., Yang, W., Astrom, S., Bystrom, A.S., and Hinnebusch, A.G. \(1995\). Modulation of tRNA\(iMet\), eIF-2, and eIF-2B expression shows that GCN4 translation is inversely coupled to the level of eIF-2.GTP.Met-tRNA\(iMet\) ternary complexes. *Mol Cell Biol* 15, 6351-6363.](#)

[Dey, M., Trieselmann, B., Locke, E.G., Lu, J., Cao, C., Dar, A.C., Krishnamoorthy, T., Dong, J., Sicheri, F., and Dever, T.E. \(2005\). PKR and GCN2 kinases and guanine nucleotide exchange factor eukaryotic translation initiation factor 2B \(eIF2B\) recognize overlapping surfaces on eIF2alpha. *Mol Cell Biol* 25, 3063-3075.](#)

[Diaconu, M., Kothe, U., Schlunzen, F., Fischer, N., Harms, J.M., Tonevitsky, A.G., Stark, H., Rodnina, M.V., and Wahl, M.C. \(2005\). Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. *Cell* 121, 991-1004.](#)

[Diedrich, G., Spahn, C.M., Stelzl, U., Schafer, M.A., Wooten, T., Bochkariov, D.E., Cooperman, B.S., Traut, R.R., and Nierhaus, K.H. \(2000\). Ribosomal protein L2 is involved in the association of the ribosomal subunits, tRNA binding to A and P sites and peptidyl transfer. *EMBO J* 19, 5241-5250.](#)

[Dmitriev, S.E., Terenin, I.M., Dunaevsky, Y.E., Merrick, W.C., and Shatsky, I.N. \(2003\). Assembly of 48S translation initiation complexes from purified components with mRNAs that have some base pairing within their 5' untranslated regions. *Mol Cell Biol* 23, 8925-8933.](#)

[Doherty, E.A., Batey, R.T., Masquida, B., and Doudna, J.A. \(2001\). A universal mode of helix packing in RNA. *Nat Struct Biol* 8, 339-343.](#)

[Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J., and Hinnebusch, A.G. \(2000\). Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. *Mol Cell* 6, 269-279.](#)

[Dong, Z., Liu, L.H., Han, B., Pincheira, R., and Zhang, J.T. \(2004\). Role of eIF3 p170 in controlling synthesis of ribonucleotide reductase M2 and cell growth. *Oncogene* 23, 3790-3801.](#)

[Dong, Z., and Zhang, J.T. \(2003\). EIF3 p170, a mediator of mimosine effect on protein synthesis and cell cycle progression. *Mol Biol Cell* 14, 3942-3951.](#)

[Donze, O., Jagus, R., Koromilas, A.E., Hershey, J.W., and Sonenberg, N. \(1995\). Abrogation of translation initiation factor eIF-2 phosphorylation causes malignant transformation of NIH 3T3 cells. *EMBO J* 14, 3828-3834.](#)

[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 40 S ribosomal subunit. *J Biol Chem* 282, 8165-8174.](#)

[Endo, Y., Mitsui, K., Motizuki, M., and Tsurugi, K. \(1987\). The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins. *J Biol Chem* 262, 5908-5912.](#)

- [Endo, Y., and Wool, I.G. \(1982\). The site of action of alpha-sarcin on eukaryotic ribosomes. The sequence at the alpha-sarcin cleavage site in 28 S ribosomal ribonucleic acid. J Biol Chem 257, 9054-9060.](#)
- [Erlacher, M.D., Lang, K., Shankaran, N., Wotzel, B., Huttenhofer, A., Micura, R., Mankin, A.S., and Polacek, N. \(2005\). Chemical engineering of the peptidyl transferase center reveals an important role of the 2'-hydroxyl group of A2451. Nucleic Acids Res 33, 1618-1627.](#)
- [Evans, D.R., Rasmussen, C., Hanic-Joyce, P.J., Johnston, G.C., Singer, R.A., and Barnes, C.A. \(1995\). Mutational analysis of the Prt1 protein subunit of yeast translation initiation factor 3. Mol Cell Biol 15, 4525-4535.](#)
- [Fekete, C.A., Applefield, D.J., Blakely, S.A., Shirokikh, N., Pestova, T., Lorsch, J.R., and Hinnebusch, A.G. \(2005\). The eIF1A C-terminal domain promotes initiation complex assembly, scanning and AUG selection *in vivo*. EMBO J 24, 3588-3601.](#)
- [Fisher, R.G., and Sweet, R.M. \(1980\). Treatment of diffraction data from crystals twinned by merohedry. Acta Cryst. A 36, 755-760.](#)
- [Fleming, K., Ghuman, J., Yuan, X., Simpson, P., Szendroi, A., Matthews, S., and Curry, S. \(2003\). Solution structure and RNA interactions of the RNA recognition motif from eukaryotic translation initiation factor 4B. Biochemistry 42, 8966-8975.](#)
- [Fletcher, C.M., Pestova, T.V., Hellen, C.U., and Wagner, G. \(1999\). Structure and interactions of the translation initiation factor eIF1. EMBO J 18, 2631-2637.](#)
- [Fontaine-Rodriguez, E.C., Taylor, T.J., Olesky, M., and Knipe, D.M. \(2004\). Proteomics of herpes simplex virus infected cell protein 27: association with translation initiation factors. Virology 330, 487-492.](#)
- [Frank, J., and Agrawal, R.K. \(2000\). A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature 406, 318-322.](#)
- [Frank, J., Gao, H., Sengupta, J., Gao, N., and Taylor, D.J. \(2007\). The process of mRNA-tRNA translocation. Proc Natl Acad Sci U S A 104, 19671-19678.](#)
- [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 40 S ribosomal subunits *in vitro*. J Biol Chem 279, 8946-8956.](#)
- [Freienstein, C., and Blobel, G. \(1975\). Nonribosomal proteins associated with eukaryotic native small ribosomal subunits. Proc Natl Acad Sci U S A 72, 3392-3396.](#)
- [Frolova, L., Le Goff, X., Rasmussen, H.H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.L., Celis, J.E., Philippe, M., and et al. \(1994\). A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. Nature 372, 701-703.](#)
- [Fukuchi-Shimogori, T., Ishii, I., Kashiwagi, K., Mashiba, H., Ekimoto, H., and Igarashi, K. \(1997\). Malignant transformation by overproduction of translation initiation factor eIF4G. Cancer Res 57, 5041-5044.](#)
- [Futterer, J., and Hohn, T. \(1991\). Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. EMBO J 10, 3887-3896.](#)
- [Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J.P. \(1987\). Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett 224, 149-155.](#)
- [Garcia, C., Fortier, P.L., Blanquet, S., Lallemand, J.Y., and Dardel, F. \(1995\). Solution structure of the ribosome-binding domain of *E. coli* translation initiation factor IF3. Homology with the U1A protein of the eukaryotic spliceosome. J Mol Biol 254, 247-259.](#)
- [Gilbert, R.J., Gordiyenko, Y., von der Haar, T., Sonnen, A.F., Hofmann, G., Nardelli, M., Stuart, D.I., and McCarthy, J.E. \(2007a\). Reconfiguration of yeast 40S ribosomal subunit domains by the translation initiation multifactor complex. Proc Natl Acad Sci U S A 104, 5788-5793.](#)

[Gilbert, W.V., Zhou, K., Butler, T.K., and Doudna, J.A. \(2007b\). Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* 317, 1224-1227.](#)

[Gingras, A.C., Raught, B., Gygi, S.P., Niedzwiecka, A., Miron, M., Burley, S.K., Polakiewicz, R.D., Wyslouch-Cieszynska, A., Aebersold, R., and Sonenberg, N. \(2001\). Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev* 15, 2852-2864.](#)

[Gold, L., Stormo, G., and Saunders, R. \(1984\). *Escherichia coli* translational initiation factor IF3: a unique case of translational regulation. *Proc Natl Acad Sci U S A* 81, 7061-7065.](#)

[Gomez, E., Mohammad, S.S., and Pavitt, G.D. \(2002\). Characterization of the minimal catalytic domain within eIF2B: the guanine-nucleotide exchange factor for translation initiation. *EMBO J* 21, 5292-5301.](#)

[Gomez, E., and Pavitt, G.D. \(2000\). Identification of domains and residues within the epsilon subunit of eukaryotic translation initiation factor 2B \(eIF2Bepsilon\) required for guanine nucleotide exchange reveals a novel activation function promoted by eIF2B complex formation. *Mol Cell Biol* 20, 3965-3976.](#)

[Gowda, S., Wu, F.C., Scholthof, H.B., and Shepherd, R.J. \(1989\). Gene VI of figwort mosaic virus \(caulimovirus group\) functions in posttranscriptional expression of genes on the full-length RNA transcript. *Proc Natl Acad Sci U S A* 86, 9203-9207.](#)

[Greenberg, J.R., Phan, L., Gu, Z., deSilva, A., Apolito, C., Sherman, F., Hinnebusch, A.G., and Goldfarb, D.S. \(1998\). Nip1p associates with 40 S ribosomes and the Prt1p subunit of eukaryotic initiation factor 3 and is required for efficient translation initiation. *J Biol Chem* 273, 23485-23494.](#)

[Grigoriadou, C., Marzi, S., Kirillov, S., Gualerzi, C.O., and Cooperman, B.S. \(2007a\). A quantitative kinetic scheme for 70 S translation initiation complex formation. *J Mol Biol* 373, 562-572.](#)

[Grigoriadou, C., Marzi, S., Pan, D., Gualerzi, C.O., and Cooperman, B.S. \(2007b\). The translational fidelity function of IF3 during transition from the 30 S initiation complex to the 70 S initiation complex. *J Mol Biol* 373, 551-561.](#)

[Groft, C.M., and Burley, S.K. \(2002\). Recognition of eIF4G by rotavirus NSP3 reveals a basis for mRNA circularization. *Mol Cell* 9, 1273-1283.](#)

[Gross, J.D., Moerke, N.J., von der Haar, T., Lugovskoy, A.A., Sachs, A.B., McCarthy, J.E., and Wagner, G. \(2003\). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739-750.](#)

[Grunberg-Manago, M., Dessen, P., Pantaloni, D., Godefroy-Colburn, T., Wolfe, A.D., and Dondon, J. \(1975\). Light-scattering studies showing the effect of initiation factors on the reversible dissociation of *Escherichia coli* ribosomes. *J Mol Biol* 94, 461-478.](#)

[Gualerzi, C., Risuleo, G., and Pon, C. \(1979\). Mechanism of the spontaneous and initiation factor 3-induced dissociation of 30 S.aminoacyl-tRNA.polynucleotide ternary complexes. *J Biol Chem* 254, 44-49.](#)

[Gualerzi, C., Risuleo, G., and Pon, C.L. \(1977\). Initial rate kinetic analysis of the mechanism of initiation complex formation and the role of initiation factor IF-3. *Biochemistry* 16, 1684-1689.](#)

[Guenneugues, M., Caserta, E., Brandi, L., Spurio, R., Meunier, S., Pon, C.L., Boelens, R., and Gualerzi, C.O. \(2000\). Mapping the fMet-tRNA\(f\)\(Met\) binding site of initiation factor IF2. *EMBO J* 19, 5233-5240.](#)

[Guillon, J.M., Heiss, S., Soutourina, J., Mechulam, Y., Laalami, S., Grunberg-Manago, M., and Blanquet, S. \(1996\). Interplay of methionine tRNAs with translation elongation factor Tu and translation initiation factor 2 in *Escherichia coli*. *J Biol Chem* 271, 22321-22325.](#)

[Guillon, J.M., Mechulam, Y., Blanquet, S., and Fayat, G. \(1993\). Importance of formylability and anticodon stem sequence to give a tRNA\(Met\) an initiator identity in *Escherichia coli*. *J Bacteriol* 175, 4507-4514.](#)

[Guillon, J.M., Mechulam, Y., Schmitter, J.M., Blanquet, S., and Fayat, G. \(1992a\). Disruption of the gene for Met-tRNA\(f\)\(Met\) formyltransferase severely impairs growth of *Escherichia coli*. *J Bacteriol* 174, 4294-4301.](#)

[Guillon, J.M., Meinel, T., Mechulam, Y., Lazennec, C., Blanquet, S., and Fayat, G. \(1992b\). Nucleotides of tRNA governing the specificity of *Escherichia coli* methionyl-tRNA\(fMet\) formyltransferase. J Mol Biol 224, 359-367.](#)

[Guillon, L., Schmitt, E., Blanquet, S., and Mechulam, Y. \(2005\). Initiator tRNA binding by e/aIF5B, the eukaryotic/archaeal homologue of bacterial initiation factor IF2. Biochemistry 44, 15594-15601.](#)

[Guo, J., Hui, D.J., Merrick, W.C., and Sen, G.C. \(2000\). A new pathway of translational regulation mediated by eukaryotic initiation factor 3. EMBO J 19, 6891-6899.](#)

[Gutierrez, P., Coillet-Matillon, S., Arrowsmith, C., and Gehring, K. \(2002\). Zinc is required for structural stability of the C-terminus of archaeal translation initiation factor aIF2beta. FEBS Lett 517, 155-158.](#)

[Haghighat, A., Mader, S., Pause, A., and Sonenberg, N. \(1995\). Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. EMBO J 14, 5701-5709.](#)

[Hanachi, P., Hershey, J.W., and Vornlocher, H.P. \(1999\). Characterization of the p33 subunit of eukaryotic translation initiation factor-3 from *Saccharomyces cerevisiae*. J Biol Chem 274, 8546-8553.](#)

[Hannig, E.M., Cigan, A.M., Freeman, B.A., and Kinzy, T.G. \(1993\). GCD11, a negative regulator of *GCN4* expression, encodes the gamma subunit of eIF-2 in *Saccharomyces cerevisiae*. Mol Cell Biol 13, 506-520.](#)

[Harris, J.K., Kelley, S.T., Spiegelman, G.B., and Pace, N.R. \(2003\). The genetic core of the universal ancestor. Genome Res 13, 407-412.](#)

[Hartwell, L.H., and McLaughlin, C.S. \(1969\). A mutant of yeast apparently defective in the initiation of protein synthesis. Proc Natl Acad Sci U S A 62, 468-474.](#)

[Hartz, D., Binkley, J., Hollingsworth, T., and Gold, L. \(1990\). Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by *Escherichia coli* IF3. Genes Dev 4, 1790-1800.](#)

[Hasek, J., Kovarik, P., Valasek, L., Malinska, K., Schneider, J., Kohlwein, S.D., and Ruis, H. \(2000\). Rpg1p, the subunit of the *Saccharomyces cerevisiae* eIF3 core complex, is a microtubule-interacting protein. Cell Motil Cytoskeleton 45, 235-246.](#)

[Hashimoto, N.N., Carnevalli, L.S., and Castilho, B.A. \(2002\). Translation initiation at non-AUG codons mediated by weakened association of eukaryotic initiation factor \(eIF\) 2 subunits. Biochem J 367, 359-368.](#)

[He, H., von der Haar, T., Singh, C.R., Li, M., Li, B., Hinnebusch, A.G., McCarthy, J.E., and Asano, K. \(2003\). The yeast eukaryotic initiation factor 4G \(eIF4G\) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. Mol Cell Biol 23, 5431-5445.](#)

[Held, D., Yaeger, K., and Novy, R. \(2004\). pCOLADuetTM-1, a lower-copy, KanR vector option for coexpression in *E. coli*. inNovations 19, 3.](#)

[Herbert, T.P., Brierley, I., and Brown, T.D. \(1997\). Identification of a protein linked to the genomic and subgenomic mRNAs of feline calicivirus and its role in translation. J Gen Virol 78 \(Pt 5\), 1033-1040.](#)

[Hershey, P.E., McWhirter, S.M., Gross, J.D., Wagner, G., Alber, T., and Sachs, A.B. \(1999\). The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. J Biol Chem 274, 21297-21304.](#)

[Heywood, S.M., Kennedy, D.S., and Bester, A.J. \(1974\). Separation of specific initiation factors involved in the translation of myosin and myoglobin messenger RNAs and the isolation of a new RNA involved in translation. Proc Natl Acad Sci U S A 71, 2428-2431.](#)

[Hinnebusch, A.G. \(1997\). Translational regulation of yeast *GCN4*. A window on factors that control initiator-trna binding to the ribosome. J Biol Chem 272, 21661-21664.](#)

[Holyoak, T., Fenn, T.D., Wilson, M.A., Moulin, A.G., Ringe, D., and Petsko, G.A. \(2003\). Malonate: a versatile cryoprotectant and stabilizing solution for salt-grown macromolecular crystals. Acta Crystallogr D Biol Crystallogr 59, 2356-2358.](#)

[Hoshino, S., Imai, M., Kobayashi, T., Uchida, N., and Katada, T. \(1999\). The eukaryotic polypeptide chain releasing factor \(eRF3/GSPT\) carrying the translation termination signal to the 3'-Poly\(A\) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. J Biol Chem 274, 16677-16680.](#)

[Howe, J.G., and Hershey, J.W. \(1983\). Initiation factor and ribosome levels are coordinately controlled in *Escherichia coli* growing at different rates. J Biol Chem 258, 1954-1959.](#)

[Huang, H.K., Yoon, H., Hannig, E.M., and Donahue, T.F. \(1997\). GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. Genes Dev 11, 2396-2413.](#)

[Hui, D.J., Terenzi, F., Merrick, W.C., and Sen, G.C. \(2005\). Mouse p56 blocks a distinct function of eukaryotic initiation factor 3 in translation initiation. J Biol Chem 280, 3433-3440.](#)

[Iacono, M., Mignone, F., and Pesole, G. \(2005\). uAUG and uORFs in human and rodent 5'untranslated mRNAs. Gene 349, 97-105.](#)

[Igarashi, N., Moriyama, H., Fujiwara, T., Fukumori, Y., and Tanaka, N. \(1997\). The 2.8 Å structure of hydroxylamine oxidoreductase from a nitrifying chemoautotrophic bacterium, *Nitrosomonas europaea*. Nat Struct Biol 4, 276-284.](#)

[Imataka, H., Gradi, A., and Sonenberg, N. \(1998\). A newly identified N-terminal amino acid sequence of human eIF4G binds poly\(A\)-binding protein and functions in poly\(A\)-dependent translation. EMBO J 17, 7480-7489.](#)

[Ito, T., Marintchev, A., and Wagner, G. \(2004\). Solution structure of human initiation factor eIF2alpha reveals homology to the elongation factor eEF1B. Structure 12, 1693-1704.](#)

[Jaramillo, M., Dever, T.E., Merrick, W.C., and Sonenberg, N. \(1991\). RNA unwinding in translation: assembly of helicase complex intermediates comprising eukaryotic initiation factors eIF-4F and eIF-4B. Mol Cell Biol 11, 5992-5997.](#)

[Jivotovskaya, A.V., Valasek, L., Hinnebusch, A.G., and Nielsen, K.H. \(2006\). Eukaryotic translation initiation factor 3 \(eIF3\) and eIF2 can promote mRNA binding to 40S subunits independently of eIF4G in yeast. Mol Cell Biol 26, 1355-1372.](#)

[Joseph, P., Lei, Y.X., Whong, W.Z., and Ong, T.M. \(2002\). Molecular cloning and functional analysis of a novel cadmium-responsive proto-oncogene. Cancer Res 62, 703-707.](#)

[Kabsch, W. \(1988\). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. J. Appl. Cryst. 21, 916-924.](#)

[Kaczanowska, M., and Ryden-Aulin, M. \(2007\). Ribosome biogenesis and the translation process in *Escherichia coli*. Microbiol Mol Biol Rev 71, 477-494.](#)

[Kapp, L.D., and Lorsch, J.R. \(2004a\). GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. J Mol Biol 335, 923-936.](#)

[Kapp, L.D., and Lorsch, J.R. \(2004b\). The molecular mechanics of eukaryotic translation. Annu Rev Biochem 73, 657-704.](#)

[Karim, M.M., Svitkin, Y.V., Kahvejian, A., De Crescenzo, G., Costa-Mattioli, M., and Sonenberg, N. \(2006\). A mechanism of translational repression by competition of Paip2 with eIF4G for poly\(A\) binding protein \(PABP\) binding. Proc Natl Acad Sci U S A 103, 9494-9499.](#)

[Karimi, R., Pavlov, M.Y., Buckingham, R.H., and Ehrenberg, M. \(1999\). Novel roles for classical factors at the interface between translation termination and initiation. Mol Cell 3, 601-609.](#)

[Kessler, S.H., and Sachs, A.B. \(1998\). RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. Mol Cell Biol 18, 51-57.](#)

[Khaleghpour, K., Svitkin, Y.V., Craig, A.W., DeMaria, C.T., Deo, R.C., Burley, S.K., and Sonenberg, N. \(2001\). Translational repression by a novel partner of human poly\(A\) binding protein, Paip2. Mol Cell 7, 205-216.](#)

[Kiesewetter, S., Ott, G., and Sprinzl, M. \(1990\). The role of modified purine 64 in initiator/elongator discrimination of tRNA\(iMet\) from yeast and wheat germ. *Nucleic Acids Res* 18, 4677-4682.](#)

[Kim, T.H., Kim, B.H., Yahalom, A., Chamovitz, D.A., and von Arnim, A.G. \(2004\). Translational regulation via 5' mRNA leader sequences revealed by mutational analysis of the Arabidopsis translation initiation factor subunit eIF3h. *Plant Cell* 16, 3341-3356.](#)

[Klein, D.J., Moore, P.B., and Steitz, T.A. \(2004\). The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* 10, 1366-1379.](#)

[Kolupaeva, V.G., Unbehaun, A., Lomakin, I.B., Hellen, C.U., and Pestova, T.V. \(2005\). Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. *RNA* 11, 470-486.](#)

[Korneeva, N.L., Lamphear, B.J., Hennigan, F.L., and Rhoads, R.E. \(2000\). Mutually cooperative binding of eukaryotic translation initiation factor \(eIF\) 3 and eIF4A to human eIF4G-1. *J Biol Chem* 275, 41369-41376.](#)

[Korostelev, A., Trakhanov, S., Laurberg, M., and Noller, H.F. \(2006\). Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell* 126, 1065-1077.](#)

[Kozak, M. \(1991\). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem* 266, 19867-19870.](#)

[Kozak, M. \(2002\). Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299, 1-34.](#)

[Kozak, M. \(2005\). A second look at cellular mRNA sequences said to function as internal ribosome entry sites. *Nucleic Acids Res* 33, 6593-6602.](#)

[Kozlov, G., Trempe, J.F., Khaleghpour, K., Kahvejian, A., Ekiel, I., and Gehring, K. \(2001\). Structure and function of the C-terminal PABC domain of human poly\(A\)-binding protein. *Proc Natl Acad Sci U S A* 98, 4409-4413.](#)

[Krishnamoorthy, T., Pavitt, G.D., Zhang, F., Dever, T.E., and Hinnebusch, A.G. \(2001\). Tight binding of the phosphorylated alpha subunit of initiation factor 2 \(eIF2alpha\) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. *Mol Cell Biol* 21, 5018-5030.](#)

[Kyrpides, N.C., and Woese, C.R. \(1998a\). Archaeal translation initiation revisited: the initiation factor 2 and eukaryotic initiation factor 2B alpha-beta-delta subunit families. *Proc Natl Acad Sci U S A* 95, 3726-3730.](#)

[Kyrpides, N.C., and Woese, C.R. \(1998b\). Universally conserved translation initiation factors. *Proc Natl Acad Sci U S A* 95, 224-228.](#)

[La Fortelle, E., and Bricogne, G. \(1997\). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* 276, 472-494.](#)

[La Teana, A., Gualerzi, C.O., and Brimacombe, R. \(1995\). From stand-by to decoding site. Adjustment of the mRNA on the 30S ribosomal subunit under the influence of the initiation factors. *RNA* 1, 772-782.](#)

[Laemmli, U.K., Molbert, E., Showe, M., and Kellenberger, E. \(1970\). Form-determining function of the genes required for the assembly of the head of bacteriophage T4. *J Mol Biol* 49, 99-113.](#)

[Lamphear, B.J., Kirchweger, R., Skern, T., and Rhoads, R.E. \(1995\). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G \(eIF4G\) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J Biol Chem* 270, 21975-21983.](#)

[Lancaster, L., and Noller, H.F. \(2005\). Involvement of 16S rRNA nucleotides G1338 and A1339 in discrimination of initiator tRNA. *Mol Cell* 20, 623-632.](#)

[Le, S.Y., and Maizel, J.V., Jr. \(1997\). A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. *Nucleic Acids Res* 25, 362-369.](#)

[Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., and Poch, O. \(2002\). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Res* 30, 5382-5390.](#)

[Lee, J.H., Choi, S.K., Roll-Mecak, A., Burley, S.K., and Dever, T.E. \(1999\). Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2. *Proc Natl Acad Sci U S A* 96, 4342-4347.](#)

[Lee, J.H., Pestova, T.V., Shin, B.S., Cao, C., Choi, S.K., and Dever, T.E. \(2002\). Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *Proc Natl Acad Sci U S A* 99, 16689-16694.](#)

[Lei, Y.X., Joseph, P., and Ong, T.M. \(2002\). Antisense inhibition of translation initiation factor 3 reverses its oncogenic potential. *Teratog Carcinog Mutagen* 22, 403-409.](#)

[Lewis, J.D., and Izaurralde, E. \(1997\). The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem* 247, 461-469.](#)

[Li, S., Kumar, N.V., Varshney, U., and RajBhandary, U.L. \(1996\). Important role of the amino acid attached to tRNA in formylation and in initiation of protein synthesis in *Escherichia coli*. *J Biol Chem* 271, 1022-1028.](#)

[Li, W., and Hoffman, D.W. \(2001\). Structure and dynamics of translation initiation factor aIF-1A from the archaeon *Methanococcus jannaschii* determined by NMR spectroscopy. *Protein Sci* 10, 2426-2438.](#)

[Lin, L., Holbro, T., Alonso, G., Gerosa, D., and Burger, M.M. \(2001\). Molecular interaction between human tumor marker protein p150, the largest subunit of eIF3, and intermediate filament protein K7. *J Cell Biochem* 80, 483-490.](#)

[Liu, Z., Dong, Z., Yang, Z., Chen, Q., Pan, Y., Yang, Y., Cui, P., Zhang, X., and Zhang, J.T. \(2007\). Role of eIF3a \(eIF3 p170\) in intestinal cell differentiation and its association with early development. *Differentiation* 75, 652-661.](#)

[Lockwood, A.H., Chakraborty, P.R., and Maitra, U. \(1971\). A complex between initiation factor IF2, guanosine triphosphate, and fMet-tRNA: an intermediate in initiation complex formation. *Proc Natl Acad Sci U S A* 68, 3122-3126.](#)

[Lomakin, I.B., Kolupaeva, V.G., Marintchev, A., Wagner, G., and Pestova, T.V. \(2003\). Position of eukaryotic initiation factor eIF1 on the 40S ribosomal subunit determined by directed hydroxyl radical probing. *Genes Dev* 17, 2786-2797.](#)

[Lomakin, I.B., Shirokikh, N.E., Yusupov, M.M., Hellen, C.U., and Pestova, T.V. \(2006\). The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH. *EMBO J* 25, 196-210.](#)

[Londei, P. \(2005\). Evolution of translational initiation: new insights from the archaea. *FEMS Microbiol Rev* 29, 185-200.](#)

[Lorsch, J.R., and Herschlag, D. \(1998a\). The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry* 37, 2180-2193.](#)

[Lorsch, J.R., and Herschlag, D. \(1998b\). The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochemistry* 37, 2194-2206.](#)

[Maag, D., Algire, M.A., and Lorsch, J.R. \(2006\). Communication between eukaryotic translation initiation factors 5 and 1A within the ribosomal pre-initiation complex plays a role in start site selection. *J Mol Biol* 356, 724-737.](#)

[Maag, D., Fekete, C.A., Gryczynski, Z., and Lorsch, J.R. \(2005\). A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. *Mol Cell* 17, 265-275.](#)

[Mader, S., Lee, H., Pause, A., and Sonenberg, N. \(1995\). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol Cell Biol* 15, 4990-4997.](#)

[Majumdar, A., Bose, K.K., and Gupta, N.K. \(1976\). Specific binding of Escherichia coli chain Initiation factor 2 to fMet-tRna^fMet. J Biol Chem 251, 137-140.](#)

[Majumdar, R., Bandyopadhyay, A., and Maitra, U. \(2003\). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. J Biol Chem 278, 6580-6587.](#)

[Majumdar, R., and Maitra, U. \(2005\). Regulation of GTP hydrolysis prior to ribosomal AUG selection during eukaryotic translation initiation. EMBO J 24, 3737-3746.](#)

[Marchetti, A., Buttitta, F., Miyazaki, S., Gallahan, D., Smith, G.H., and Callahan, R. \(1995\). Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. J Virol 69, 1932-1938.](#)

[Marchetti, A., Buttitta, F., Pellegrini, S., Bertacca, G., and Callahan, R. \(2001\). Reduced expression of INT-6/eIF3-p48 in human tumors. Int J Oncol 18, 175-179.](#)

[Marck, C. \(1988\). 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res 16, 1829-1836.](#)

[Marck, C., and Grosjean, H. \(2002\). tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. RNA 8, 1189-1232.](#)

[Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. \(1997\). Cocystal structure of the messenger RNA 5' cap-binding protein \(eIF4E\) bound to 7-methyl-GDP. Cell 89, 951-961.](#)

[Marcotrigiano, J., Lomakin, I.B., Sonenberg, N., Pestova, T.V., Hellen, C.U., and Burley, S.K. \(2001\). A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. Mol Cell 7, 193-203.](#)

[Marintchev, A., Kolupaeva, V.G., Pestova, T.V., and Wagner, G. \(2003\). Mapping the binding interface between human eukaryotic initiation factors 1A and 5B: a new interaction between old partners. Proc Natl Acad Sci U S A 100, 1535-1540.](#)

[Matsuda, S., Katsumata, R., Okuda, T., Yamamoto, T., Miyazaki, K., Senga, T., Machida, K., Thant, A.A., Nakatsugawa, S., and Hamaguchi, M. \(2000\). Molecular cloning and characterization of human MAWD, a novel protein containing WD-40 repeats frequently overexpressed in breast cancer. Cancer Res 60, 13-17.](#)

[Matsuo, H., Li, H., McGuire, A.M., Fletcher, C.M., Gingras, A.C., Sonenberg, N., and Wagner, G. \(1997\). Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nat Struct Biol 4, 717-724.](#)

[Mayeur, G.L., Fraser, C.S., Peiretti, F., Block, K.L., and Hershey, J.W. \(2003\). Characterization of eIF3k: a newly discovered subunit of mammalian translation initiation factor eIF3. Eur J Biochem 270, 4133-4139.](#)

[Mayeur, G.L., and Hershey, J.W. \(2002\). Malignant transformation by the eukaryotic translation initiation factor 3 subunit p48 \(eIF3e\). FEBS Lett 514, 49-54.](#)

[McFerrin, M.B., and Snell, E.H. \(2002\). The development and application of a method to quantify the quality of cryoprotectant solutions using standard area-detector X-ray images. Journal of Applied Crystallography 35, 538-545.](#)

[Mears, J.A., Cannone, J.J., Stagg, S.M., Gutell, R.R., Agrawal, R.K., and Harvey, S.C. \(2002\). Modeling a minimal ribosome based on comparative sequence analysis. J Mol Biol 321, 215-234.](#)

[Mechulam, Y., Guillon, L., Yatime, L., Blanquet, S., and Schmitt, E. \(2007\). Protection-based assays to measure aminoacyl-tRNA binding to translation initiation factors. Methods Enzymol 430, 265-281.](#)

[Meinzel, T., Mechulam, Y., and Fayat, G. \(1988\). Fast purification of a functional elongator tRNA^{Met} expressed from a synthetic gene *in vivo*. Nucleic Acids Res 16, 8095-8096.](#)

[Meinzel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S., and Fayat, G. \(1991\). Selection of suppressor methionyl-tRNA synthetases: mapping the tRNA anticodon binding site. *Proc Natl Acad Sci U S A* 88, 291-295.](#)

[Meinzel, T., Sacerdot, C., Graffe, M., Blanquet, S., and Springer, M. \(1999\). Discrimination by *Escherichia coli* initiation factor IF3 against initiation on non-canonical codons relies on complementarity rules. *J Mol Biol* 290, 825-837.](#)

[Mellot, P., Mechulam, Y., Le Corre, D., Blanquet, S., and Fayat, G. \(1989\). Identification of an amino acid region supporting specific methionyl-tRNA synthetase: tRNA recognition. *J Mol Biol* 208, 429-443.](#)

[Merrick, W.C., Kemper, W.M., and Anderson, W.F. \(1975\). Purification and characterization of homogeneous initiation factor M2A from rabbit reticulocytes. *J Biol Chem* 250, 5556-5562.](#)

[Methot, N., Song, M.S., and Sonenberg, N. \(1996\). 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, 5328-5334.](#)

Miller, T.Y., He, X.M., and Carter, D.C. (1992). A comparison between protein crystals grown with vapor diffusion methods in microgravity and protein crystals using a gel liquid-liquid diffusion ground-based method. *J. Cryst. Growth* 122, 306-309.

[Minich, W.B., Balasta, M.L., Goss, D.J., and Rhoads, R.E. \(1994\). Chromatographic resolution of *in vivo* phosphorylated and nonphosphorylated eukaryotic translation initiation factor eIF-4E: increased cap affinity of the phosphorylated form. *Proc Natl Acad Sci U S A* 91, 7668-7672.](#)

[Moll, I., Grill, S., Gualerzi, C.O., and Blasi, U. \(2002\). Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol Microbiol* 43, 239-246.](#)

[Moll, I., Hirokawa, G., Kiel, M.C., Kaji, A., and Blasi, U. \(2004\). Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. *Nucleic Acids Res* 32, 3354-3363.](#)

[Morgan, D.G., Menetret, J.F., Neuhof, A., Rapoport, T.A., and Akey, C.W. \(2002\). Structure of the mammalian ribosome-channel complex at 17Å resolution. *J Mol Biol* 324, 871-886.](#)

[Morgan, D.G., Menetret, J.F., Radermacher, M., Neuhof, A., Akey, I.V., Rapoport, T.A., and Akey, C.W. \(2000\). A comparison of the yeast and rabbit 80 S ribosome reveals the topology of the nascent chain exit tunnel, inter-subunit bridges and mammalian rRNA expansion segments. *J Mol Biol* 301, 301-321.](#)

[Myasnikov, A.G., Marzi, S., Simonetti, A., Giuliodori, A.M., Gualerzi, C.O., Yusupova, G., Yusupov, M., and Klaholz, B.P. \(2005\). Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. *Nat Struct Mol Biol* 12, 1145-1149.](#)

[Nakai, Y., Shiratsuchi, A., Manaka, J., Nakayama, H., Takio, K., Zhang, J.T., Suganuma, T., and Nakanishi, Y. \(2005\). Externalization and recognition by macrophages of large subunit of eukaryotic translation initiation factor 3 in apoptotic cells. *Exp Cell Res* 309, 137-148.](#)

[Naranda, T., Kainuma, M., MacMillan, S.E., and Hershey, J.W. \(1997\). The 39-kilodalton subunit of eukaryotic translation initiation factor 3 is essential for the complex's integrity and for cell viability in *Saccharomyces cerevisiae*. *Mol Cell Biol* 17, 145-153.](#)

[Naranda, T., MacMillan, S.E., and Hershey, J.W. \(1994\). Purified yeast translational initiation factor eIF-3 is an RNA-binding protein complex that contains the PRT1 protein. *J Biol Chem* 269, 32286-32292.](#)

[Nielsen, K.H., Szamecz, B., Valasek, L., Jivotovskaya, A., Shin, B.S., and Hinnebusch, A.G. \(2004\). Functions of eIF3 downstream of 48S assembly impact AUG recognition and *GCN4* translational control. *EMBO J* 23, 1166-1177.](#)

[Nielsen, K.H., Valasek, L., Sykes, C., Jivotovskaya, A., and Hinnebusch, A.G. \(2006\). Interaction of the RNP1 motif in PRT1 with HCR1 promotes 40S binding of eukaryotic initiation factor 3 in yeast. *Mol Cell Biol* 26, 2984-2998.](#)

[Nika, J., Rippel, S., and Hannig, E.M. \(2001\). Biochemical analysis of the eIF2beta gamma complex reveals a structural function for eIF2alpha in catalyzed nucleotide exchange. *J Biol Chem* 276, 1051-1056.](#)

- [Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B., and Steitz, T.A. \(2001\). RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. Proc Natl Acad Sci U S A 98, 4899-4903.](#)
- Novy, R., Drott, D., Yaeger, K., and Mierendorf, R. (2001). Overcoming the codon bias of *E. coli* for enhanced protein expression. *In* *Novations* 12, 3.
- [Nupponen, N.N., Porkka, K., Kakkola, L., Tanner, M., Persson, K., Borg, A., Isola, J., and Visakorpi, T. \(1999\). Amplification and overexpression of p40 subunit of eukaryotic translation initiation factor 3 in breast and prostate cancer. Am J Pathol 154, 1777-1783.](#)
- [O'Connor, M., Gregory, S.T., Rajbhandary, U.L., and Dahlberg, A.E. \(2001\). Altered discrimination of start codons and initiator tRNAs by mutant initiation factor 3. RNA 7, 969-978.](#)
- [O'Donnell, S.M., and Janssen, G.R. \(2002\). Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in *Escherichia coli*. J Bacteriol 184, 6730-6733.](#)
- [Oberer, M., Marintchev, A., and Wagner, G. \(2005\). Structural basis for the enhancement of eIF4A helicase activity by eIF4G. Genes Dev 19, 2212-2223.](#)
- [Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. \(2001\). Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science 292, 897-902.](#)
- [Okamoto, H., Yasui, K., Zhao, C., Arai, S., and Inazawa, J. \(2003\). PTK2 and EIF3S3 genes may be amplification targets at 8q23-q24 and are associated with large hepatocellular carcinomas. Hepatology 38, 1242-1249.](#)
- [Olsen, D.S., Savner, E.M., Mathew, A., Zhang, F., Krishnamoorthy, T., Phan, L., and Hinnebusch, A.G. \(2003\). Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment *in vivo*. EMBO J 22, 193-204.](#)
- [Otero, L.J., Ashe, M.P., and Sachs, A.B. \(1999\). The yeast poly\(A\)-binding protein Pab1p stimulates *in vitro* poly\(A\)-dependent and cap-dependent translation by distinct mechanisms. EMBO J 18, 3153-3163.](#)
- Otwinoski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. *Isomorphous Replacement and Anomalous Scattering* (eds. W. Wolf, P.R. Evans, and A.G.W. Leslie), 80-86.
- [Park, H.S., Browning, K.S., Hohn, T., and Ryabova, L.A. \(2004\). Eucaryotic initiation factor 4B controls eIF3-mediated ribosomal entry of viral reinitiation factor. EMBO J 23, 1381-1391.](#)
- [Park, H.S., Himmelbach, A., Browning, K.S., Hohn, T., and Ryabova, L.A. \(2001\). A plant viral "reinitiation" factor interacts with the host translational machinery. Cell 106, 723-733.](#)
- [Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., and Ramakrishnan, V. \(2007\). The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. Mol Cell 26, 41-50.](#)
- [Pause, A., Belsham, G.J., Gingras, A.C., Donze, O., Lin, T.A., Lawrence, J.C., Jr., and Sonenberg, N. \(1994\). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. Nature 371, 762-767.](#)
- [Pavitt, G.D., Ramaiah, K.V., Kimball, S.R., and Hinnebusch, A.G. \(1998\). eIF2 independently binds two distinct eIF2B subcomplexes that catalyze and regulate guanine-nucleotide exchange. Genes Dev 12, 514-526.](#)
- [Pestova, T.V., Borukhov, S.I., and Hellen, C.U. \(1998\). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. Nature 394, 854-859.](#)
- [Pestova, T.V., and Hellen, C.U. \(2000\). The structure and function of initiation factors in eukaryotic protein synthesis. Cell Mol Life Sci 57, 651-674.](#)
- [Pestova, T.V., and Hellen, C.U. \(2003\). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. Genes Dev 17, 181-186.](#)
- [Pestova, T.V., Hellen, C.U., and Shatsky, I.N. \(1996a\). Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Mol Cell Biol 16, 6859-6869.](#)

[Pestova, T.V., and Kolupaeva, V.G. \(2002\). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* 16, 2906-2922.](#)

[Pestova, T.V., Lomakin, I.B., Lee, J.H., Choi, S.K., Dever, T.E., and Hellen, C.U. \(2000\). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* 403, 332-335.](#)

[Pestova, T.V., Shatsky, I.N., and Hellen, C.U. \(1996b\). Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol* 16, 6870-6878.](#)

[Petersen, H.U., Roll, T., Grunberg-Manago, M., and Clark, B.F. \(1979\). Specific interaction of initiation factor IF2 of *E. coli* with formylmethionyl-tRNA f Met. *Biochem Biophys Res Commun* 91, 1068-1074.](#)

[Phan, L., Schoenfeld, L.W., Valasek, L., Nielsen, K.H., and Hinnebusch, A.G. \(2001\). A subcomplex of three eIF3 subunits binds eIF1 and eIF5 and stimulates ribosome binding of mRNA and tRNA\(i\)Met. *EMBO J* 20, 2954-2965.](#)

[Phan, L., Zhang, X., Asano, K., Anderson, J., Vornlocher, H.P., Greenberg, J.R., Qin, J., and Hinnebusch, A.G. \(1998\). Identification of a translation initiation factor 3 \(eIF3\) core complex, conserved in yeast and mammals, that interacts with eIF5. *Mol Cell Biol* 18, 4935-4946.](#)

[Pilipenko, E.V., Pestova, T.V., Kolupaeva, V.G., Khitrina, E.V., Poperechnaya, A.N., Agol, V.I., and Hellen, C.U. \(2000\). A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev* 14, 2028-2045.](#)

[Pincheira, R., Chen, Q., Huang, Z., and Zhang, J.T. \(2001a\). Two subcellular localizations of eIF3 p170 and its interaction with membrane-bound microfilaments: implications for alternative functions of p170. *Eur J Cell Biol* 80, 410-418.](#)

[Pincheira, R., Chen, Q., and Zhang, J.T. \(2001b\). Identification of a 170-kDa protein over-expressed in lung cancers. *Br J Cancer* 84, 1520-1527.](#)

[Pisarev, A.V., Kolupaeva, V.G., Pisareva, V.P., Merrick, W.C., Hellen, C.U., and Pestova, T.V. \(2006\). Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev* 20, 624-636.](#)

[Plewniak, F., Bianchetti, L., Brelivet, Y., Carles, A., Chalmel, F., Lecompte, O., Mochel, T., Moulinier, L., Muller, A., Muller, J., et al. \(2003\). PipeAlign: A new toolkit for protein family analysis. *Nucleic Acids Res* 31, 3829-3832.](#)

[Plewniak, F., Thompson, J.D., and Poch, O. \(2000\). Ballast: blast post-processing based on locally conserved segments. *Bioinformatics* 16, 750-759.](#)

[Pon, C.L., and Gualerzi, C.O. \(1984\). Mechanism of protein biosynthesis in prokaryotic cells. Effect of initiation factor IF1 on the initial rate of 30 S initiation complex formation. *FEBS Lett* 175, 203-207.](#)

[Pon, C.L., Paci, M., Pawlik, R.T., and Gualerzi, C.O. \(1985\). Structure-function relationship in *Escherichia coli* initiation factors. Biochemical and biophysical characterization of the interaction between IF-2 and guanosine nucleotides. *J Biol Chem* 260, 8918-8924.](#)

[Poyry, T.A., Kaminski, A., and Jackson, R.J. \(2004\). What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev* 18, 62-75.](#)

[Preiss, T., and M, W.H. \(2003\). Starting the protein synthesis machine: eukaryotic translation initiation. *Bioessays* 25, 1201-1211.](#)

[Pyrnnet, S., Imataka, H., Gingras, A.C., Fukunaga, R., Hunter, T., and Sonenberg, N. \(1999\). Human eukaryotic translation initiation factor 4G \(eIF4G\) recruits mnk1 to phosphorylate eIF4E. *EMBO J* 18, 270-279.](#)

[Ramakrishnan, V. \(2002\). Ribosome structure and the mechanism of translation. *Cell* 108, 557-572.](#)

[Ray, B.K., Lawson, T.G., Kramer, J.C., Cladaras, M.H., Grifo, J.A., Abramson, R.D., Merrick, W.C., and Thach, R.E. \(1985\). ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. *J Biol Chem* 260, 7651-7658.](#)

- [Rees, D.C., and Lipscomb, W.N. \(1980\). Structure of potato inhibitor complex of carboxypeptidase A at 5.5-Å resolution. Proc Natl Acad Sci U S A 77, 277-280.](#)
- [Rogers, G.W., Jr., Richter, N.J., Lima, W.F., and Merrick, W.C. \(2001\). Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. J Biol Chem 276, 30914-30922.](#)
- [Roll-Mecak, A., Cao, C., Dever, T.E., and Burley, S.K. \(2000\). X-Ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. Cell 103, 781-792.](#)
- [Rothe, M., Ko, Y., Albers, P., and Wernert, N. \(2000\). Eukaryotic initiation factor 3 p110 mRNA is overexpressed in testicular seminomas. Am J Pathol 157, 1597-1604.](#)
- [Rowlands, A.G., Panniers, R., and Henshaw, E.C. \(1988\). The catalytic mechanism of guanine nucleotide exchange factor action and competitive inhibition by phosphorylated eukaryotic initiation factor 2. J Biol Chem 263, 5526-5533.](#)
- [Roy, G., De Crescenzo, G., Khaleghpour, K., Kahvejian, A., O'Connor-McCourt, M., and Sonenberg, N. \(2002\). Paip1 interacts with poly\(A\) binding protein through two independent binding motifs. Mol Cell Biol 22, 3769-3782.](#)
- [Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C., and Sonenberg, N. \(1990\). Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. Mol Cell Biol 10, 1134-1144.](#)
- [Rubinson, K.A., Ladner, J.E., Tordova, M., and Gilliland, G.L. \(2000\). Cryosalts: suppression of ice formation in macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 56, 996-1001.](#)
- [Sacerdot, C., Vachon, G., Laalami, S., Morel-Deville, F., Cenatiempo, Y., and Grunberg-Manago, M. \(1992\). Both forms of translational initiation factor IF2 \(alpha and beta\) are required for maximal growth of *Escherichia coli*. Evidence for two translational initiation codons for IF2 beta. J Mol Biol 225, 67-80.](#)
- [Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F., and Petersen, G.B. \(1982\). Nucleotide sequence of bacteriophage lambda DNA. J Mol Biol 162, 729-773.](#)
- [Sanvito, F., Piatti, S., Villa, A., Bossi, M., Lucchini, G., Marchisio, P.C., and Biffo, S. \(1999\). The beta4 integrin interactor p27\(BBP/eIF6\) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly. J Cell Biol 144, 823-837.](#)
- [Saramaki, O., Willi, N., Bratt, O., Gasser, T.C., Koivisto, P., Nupponen, N.N., Bubendorf, L., and Visakorpi, T. \(2001\). Amplification of EIF3S3 gene is associated with advanced stage in prostate cancer. Am J Pathol 159, 2089-2094.](#)
- [Savelsbergh, A., Mohr, D., Wilden, B., Wintermeyer, W., and Rodnina, M.V. \(2000\). Stimulation of the GTPase activity of translation elongation factor G by ribosomal protein L7/12. J Biol Chem 275, 890-894.](#)
- [Savinainen, K.J., Linja, M.J., Saramaki, O.R., Tammela, T.L., Chang, G.T., Brinkmann, A.O., and Visakorpi, T. \(2004\). Expression and copy number analysis of TRPS1, EIF3S3 and MYC genes in breast and prostate cancer. Br J Cancer 90, 1041-1046.](#)
- [Scheper, G.C., and Proud, C.G. \(2002\). Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? Eur J Biochem 269, 5350-5359.](#)
- [Scheper, G.C., van Kollenburg, B., Hu, J., Luo, Y., Goss, D.J., and Proud, C.G. \(2002\). Phosphorylation of eukaryotic initiation factor 4E markedly reduces its affinity for capped mRNA. J Biol Chem 277, 3303-3309.](#)
- [Schevitz, R.W., Podjarny, A.D., Krishnamachari, N., Hughes, J.J., Sigler, P.B., and Sussman, J.L. \(1979\). Crystal structure of a eukaryotic initiator tRNA. Nature 278, 188-190.](#)
- [Schmitt, E., Blanquet, S., and Mechulam, Y. \(2002\). The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors. EMBO J 21, 1821-1832.](#)
- [Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. \(1998\). Crystal structure of methionyl-tRNA^{Met} transformylase complexed with the initiator formyl-methionyl-tRNA^{Met}. EMBO J 17, 6819-6826.](#)

[Schmitter, J.M., Mechulam, Y., Fayat, G., and Anselme, M. \(1986\). Rapid purification of DNA fragments by high-performance size-exclusion chromatography. J Chromatogr 378, 462-466.](#)

[Schuetz, A., Allali-Hassani, A., Martin, F., Loppnau, P., Vedadi, M., Bochkarev, A., Plotnikov, A.N., Arrowsmith, C.H., and Min, J. \(2006\). Structural basis for molecular recognition and presentation of histone H3 by WDR5. EMBO J 25, 4245-4252.](#)

[Schuler, M., Connell, S.R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., Mielke, T., Penczek, P.A., Westhof, E., and Spahn, C.M. \(2006\). Structure of the ribosome-bound cricket paralysis virus IRES RNA. Nat Struct Mol Biol 13, 1092-1096.](#)

[Schulman, L.H., and Pelka, H. \(1988\). Anticodon switching changes the identity of methionine and valine transfer RNAs. Science 242, 765-768.](#)

[Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. \(2005\). Structures of the bacterial ribosome at 3.5 Å resolution. Science 310, 827-834.](#)

[Selmer, M., Dunham, C.M., Murphy, F.V.t., Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R., and Ramakrishnan, V. \(2006\). Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313, 1935-1942.](#)

[Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. \(2006\). Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. Cell 125, 287-300.](#)

[Seong, B.L., and RajBhandary, U.L. \(1987a\). *Escherichia coli* formylmethionine tRNA: mutations in GGGCC sequence conserved in anticodon stem of initiator tRNAs affect initiation of protein synthesis and conformation of anticodon loop. Proc Natl Acad Sci U S A 84, 334-338.](#)

[Seong, B.L., and RajBhandary, U.L. \(1987b\). Mutants of *Escherichia coli* formylmethionine tRNA: a single base change enables initiator tRNA to act as an elongator *in vitro*. Proc Natl Acad Sci U S A 84, 8859-8863.](#)

[Sette, M., van Tilborg, P., Spurio, R., Kaptein, R., Paci, M., Gualerzi, C.O., and Boelens, R. \(1997\). The structure of the translational initiation factor IF1 from E.coli contains an oligomer-binding motif. EMBO J 16, 1436-1443.](#)

[Shalev, A., Valasek, L., Pise-Masison, C.A., Radonovich, M., Phan, L., Clayton, J., He, H., Brady, J.N., Hinnebusch, A.G., and Asano, K. \(2001\). *Saccharomyces cerevisiae* protein Pci8p and human protein eIF3e/Int-6 interact with the eIF3 core complex by binding to cognate eIF3b subunits. J Biol Chem 276, 34948-34957.](#)

[Shanina, N.A., Ivanov, P.A., Chudinova, E.M., Severin, F.F., and Nadezhdina, E.S. \(2001\). \[Translation initiation factor eIF3 is able to bind with microtubules in mammalian cells\]. Mol Biol \(Mosk\) 35, 638-646.](#)

[Shatkin, A.J. \(1976\). Capping of eucaryotic mRNAs. Cell 9, 645-653.](#)

[Shi, J., Feng, Y., Goulet, A.C., Vaillancourt, R.R., Sachs, N.A., Hershey, J.W., and Nelson, M.A. \(2003\). The p34cdc2-related cyclin-dependent kinase 11 interacts with the p47 subunit of eukaryotic initiation factor 3 during apoptosis. J Biol Chem 278, 5062-5071.](#)

[Shin, B.S., Maag, D., Roll-Mecak, A., Arefin, M.S., Burley, S.K., Lorsch, J.R., and Dever, T.E. \(2002\). Uncoupling of initiation factor eIF5B/IF2 GTPase and translational activities by mutations that lower ribosome affinity. Cell 111, 1015-1025.](#)

[Shine, J., and Dalgarno, L. \(1974\). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci U S A 71, 1342-1346.](#)

[Si, K., Chaudhuri, J., Chevesich, J., and Maitra, U. \(1997\). Molecular cloning and functional expression of a human cDNA encoding translation initiation factor 6. Proc Natl Acad Sci U S A 94, 14285-14290.](#)

[Si, K., and Maitra, U. \(1999\). The *Saccharomyces cerevisiae* homologue of mammalian translation initiation factor 6 does not function as a translation initiation factor. Mol Cell Biol 19, 1416-1426.](#)

[Singh, C.R., Curtis, C., Yamamoto, Y., Hall, N.S., Kruse, D.S., He, H., Hannig, E.M., and Asano, K. \(2005\). Eukaryotic translation initiation factor 5 is critical for integrity of the scanning preinitiation complex and accurate control of GCN4 translation. *Mol Cell Biol* 25, 5480-5491.](#)

[Singh, C.R., Yamamoto, Y., and Asano, K. \(2004\). Physical association of eukaryotic initiation factor \(eIF\) 5 carboxyl-terminal domain with the lysine-rich eIF2beta segment strongly enhances its binding to eIF3. *J Biol Chem* 279, 49644-49655.](#)

[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.](#)

[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.](#)

[Smith, T.F., Gaitatzes, C., Saxena, K., and Neer, E.J. \(1999\). The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci* 24, 181-185.](#)

[Sokabe, M., Yao, M., Sakai, N., Toya, S., and Tanaka, I. \(2006\). Structure of archaeal translational initiation factor 2 betagamma-GDP reveals significant conformational change of the beta-subunit and switch 1 region. *Proc Natl Acad Sci U S A* 103, 13016-13021.](#)

[Spahn, C.M., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., and Frank, J. \(2001\). Structure of the 80S ribosome from *Saccharomyces cerevisiae*--tRNA-ribosome and subunit-subunit interactions. *Cell* 107, 373-386.](#)

[Spahn, C.M., Gomez-Lorenzo, M.G., Grassucci, R.A., Jorgensen, R., Andersen, G.R., Beckmann, R., Penczek, P.A., Ballesta, J.P., and Frank, J. \(2004a\). Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J* 23, 1008-1019.](#)

[Spahn, C.M., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P., and Frank, J. \(2004b\). Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* 118, 465-475.](#)

[Spurio, R., Brandi, L., Caserta, E., Pon, C.L., Gualerzi, C.O., Misselwitz, R., Krafft, C., Welfle, K., and Welfle, H. \(2000\). The C-terminal subdomain \(IF2 C-2\) contains the entire fMet-tRNA binding site of initiation factor IF2. *J Biol Chem* 275, 2447-2454.](#)

[Storoni, L.C., McCoy, A.J., and Read, R.J. \(2004\). Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 60, 432-438.](#)

[Studier, F.W. \(2005\). Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41, 207-234.](#)

[Sundari, R.M., Stringer, E.A., Schulman, L.H., and Maitra, U. \(1976\). Interaction of bacterial initiation factor 2 with initiator tRNA. *J Biol Chem* 251, 3338-3345.](#)

[Sussman, J.K., Simons, E.L., and Simons, R.W. \(1996\). *Escherichia coli* translation initiation factor 3 discriminates the initiation codon *in vivo*. *Mol Microbiol* 21, 347-360.](#)

[Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Ishida, T., Taniguchi, T., Hasegawa, H., Terashima, A., et al. \(2002\). Crystal structures of 7-methylguanosine 5'-triphosphate \(m\(7\)GTP\)- and P\(1\)-7-methylguanosine-P\(3\)-adenosine-5',5'-triphosphate \(m\(7\)GpppA\)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochem J* 362, 539-544.](#)

[Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Sasaki, M., Taniguchi, T., Miyagawa, H., Kitamura, K., et al. \(2003\). Structural features of human initiation factor 4E, studied by X-ray crystal analyses and molecular dynamics simulations. *J Mol Biol* 328, 365-383.](#)

[Tomsic, J., Vitali, L.A., Daviter, T., Savelsbergh, A., Spurio, R., Striebeck, P., Wintermeyer, W., Rodnina, M.V., and Gualerzi, C.O. \(2000\). Late events of translation initiation in bacteria: a kinetic analysis. *EMBO J* 19, 2127-2136.](#)

[Trachsel, H., Erni, B., Schreier, M.H., and Staehelin, T. \(1977\). Initiation of mammalian protein synthesis. II. The assembly of the initiation complex with purified initiation factors. J Mol Biol 116, 755-767.](#)

[Tschochner, H., and Hurt, E. \(2003\). Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol 13, 255-263.](#)

[Udagawa, T., Shimizu, Y., and Ueda, T. \(2004\). Evidence for the translation initiation of leaderless mRNAs by the intact 70 S ribosome without its dissociation into subunits in eubacteria. J Biol Chem 279, 8539-8546.](#)

[Unbehaun, A., Borukhov, S.I., Hellen, C.U., and Pestova, T.V. \(2004\). Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. Genes Dev 18, 3078-3093.](#)

[Valasek, L., Hasek, J., Trachsel, H., Imre, E.M., and Ruis, H. \(1999\). The *Saccharomyces cerevisiae* HCR1 gene encoding a homologue of the p35 subunit of human translation initiation factor 3 \(eIF3\) is a high copy suppressor of a temperature-sensitive mutation in the Rpg1p subunit of yeast eIF3. J Biol Chem 274, 27567-27572.](#)

[Valasek, L., Mathew, A.A., Shin, B.S., Nielsen, K.H., Szamecz, B., and Hinnebusch, A.G. \(2003\). The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome *in vivo*. Genes Dev 17, 786-799.](#)

[Valasek, L., Nielsen, K.H., and Hinnebusch, A.G. \(2002\). Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation *in vivo*. EMBO J 21, 5886-5898.](#)

[Valasek, L., Nielsen, K.H., Zhang, F., Fekete, C.A., and Hinnebusch, A.G. \(2004\). Interactions of eukaryotic translation initiation factor 3 \(eIF3\) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. Mol Cell Biol 24, 9437-9455.](#)

[Valasek, L., Phan, L., Schoenfeld, L.W., Valaskova, V., and Hinnebusch, A.G. \(2001\). Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. EMBO J 20, 891-904.](#)

[Valasek, L., Trachsel, H., Hasek, J., and Ruis, H. \(1998\). Rpg1, the *Saccharomyces cerevisiae* homologue of the largest subunit of mammalian translation initiation factor 3, is required for translational activity. J Biol Chem 273, 21253-21260.](#)

[Valegard, K., van Scheltinga, A.C., Lloyd, M.D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H.J., Baldwin, J.E., Schofield, C.J., et al. \(1998\). Structure of a cephalosporin synthase. Nature 394, 805-809.](#)

[Vattem, K.M., and Wek, R.C. \(2004\). Reinitiation involving upstream ORFs regulates *ATF4* mRNA translation in mammalian cells. Proc Natl Acad Sci U S A 101, 11269-11274.](#)

[Vazquez de Aldana, C.R., Dever, T.E., and Hinnebusch, A.G. \(1993\). Mutations in the alpha subunit of eukaryotic translation initiation factor 2 \(eIF-2 alpha\) that overcome the inhibitory effect of eIF-2 alpha phosphorylation on translation initiation. Proc Natl Acad Sci U S A 90, 7215-7219.](#)

[Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P., and Bradbury, A.R. \(2007\). Plasmid incompatibility: more compatible than previously thought? Protein Eng Des Sel 20, 309-313.](#)

[Vende, P., Piron, M., Castagne, N., and Poncet, D. \(2000\). Efficient translation of rotavirus mRNA requires simultaneous interaction of NSP3 with the eukaryotic translation initiation factor eIF4G and the mRNA 3' end. J Virol 74, 7064-7071.](#)

[Verlhac, M.H., Chen, R.H., Hanachi, P., Hershey, J.W., and Derynck, R. \(1997\). Identification of partners of *TIF34*, a component of the yeast eIF3 complex, required for cell proliferation and translation initiation. EMBO J 16, 6812-6822.](#)

[Vornlocher, H.P., Hanachi, P., Ribeiro, S., and Hershey, J.W. \(1999\). A 110-kilodalton subunit of translation initiation factor eIF3 and an associated 135-kilodalton protein are encoded by the *Saccharomyces cerevisiae* TIF32 and TIF31 genes. J Biol Chem 274, 16802-16812.](#)

[Warner, J.R. \(1999\). The economics of ribosome biosynthesis in yeast. Trends Biochem Sci 24, 437-440.](#)

[Wei, Z., Xue, Y., Xu, H., and Gong, W. \(2006\). Crystal structure of the C-terminal domain of *S.cerevisiae* eIF5. J Mol Biol 359, 1-9.](#)

[Wei, Z., Zhang, P., Zhou, Z., Cheng, Z., Wan, M., and Gong, W. \(2004\). Crystal structure of human eIF3k, the first structure of eIF3 subunits. J Biol Chem 279, 34983-34990.](#)

[Weiel, J., and Hershey, J.W. \(1982\). The binding of fluorescein-labeled protein synthesis initiation factor 2 to *Escherichia coli* 30 S ribosomal subunits determined by fluorescence polarization. J Biol Chem 257, 1215-1220.](#)

[Weinger, J.S., Parnell, K.M., Dorner, S., Green, R., and Strobel, S.A. \(2004\). Substrate-assisted catalysis of peptide bond formation by the ribosome. Nat Struct Mol Biol 11, 1101-1106.](#)

[Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. \(1998\). Circularization of mRNA by eukaryotic translation initiation factors. Mol Cell 2, 135-140.](#)

[Woese, C.R., and Fox, G.E. \(1977\). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A 74, 5088-5090.](#)

[Woese, C.R., Kandler, O., and Wheelis, M.L. \(1990\). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A 87, 4576-4579.](#)

[Woese, C.R., Magrum, L.J., and Fox, G.E. \(1978\). Archaeobacteria. J Mol Evol 11, 245-251.](#)

[Wu, X.Q., and RajBhandary, U.L. \(1997\). Effect of the amino acid attached to *Escherichia coli* initiator tRNA on its affinity for the initiation factor IF2 and on the IF2 dependence of its binding to the ribosome. J Biol Chem 272, 1891-1895.](#)

[Yamamoto, Y., Singh, C.R., Marintchev, A., Hall, N.S., Hannig, E.M., Wagner, G., and Asano, K. \(2005\). The eukaryotic initiation factor \(eIF\) 5 HEAT domain mediates multifactor assembly and scanning with distinct interfaces to eIF1, eIF2, eIF3, and eIF4G. Proc Natl Acad Sci U S A 102, 16164-16169.](#)

[Yang, W., and Hinnebusch, A.G. \(1996\). Identification of a regulatory subcomplex in the guanine nucleotide exchange factor eIF2B that mediates inhibition by phosphorylated eIF2. Mol Cell Biol 16, 6603-6616.](#)

[Yatime, L., Mechulam, Y., Blanquet, S., and Schmitt, E. \(2006\). Structural switch of the gamma subunit in an archaeal aIF2 alpha gamma heterodimer. Structure 14, 119-128.](#)

[Yatime, L., Mechulam, Y., Blanquet, S., and Schmitt, E. \(2007\). Structure of an archaeal heterotrimeric initiation factor 2 reveals a nucleotide state between the GTP and the GDP states. Proc Natl Acad Sci U S A 104, 18445-18450.](#)

[Yatime, L., Schmitt, E., Blanquet, S., and Mechulam, Y. \(2004\). Functional molecular mapping of archaeal translation initiation factor 2. J Biol Chem 279, 15984-15993.](#)

[Yatime, L., Schmitt, E., Blanquet, S., and Mechulam, Y. \(2005\). Structure-function relationships of the intact aIF2alpha subunit from the archaeon *Pyrococcus abyssi*. Biochemistry 44, 8749-8756.](#)

[Yoon, H.J., and Donahue, T.F. \(1992\). The suil suppressor locus in *Saccharomyces cerevisiae* encodes a translation factor that functions during tRNA\(iMet\) recognition of the start codon. Mol Cell Biol 12, 248-260.](#)

[Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., and Noller, H.F. \(2001\). Crystal structure of the ribosome at 5.5 A resolution. Science 292, 883-896.](#)

[Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. \(2001\). The path of messenger RNA through the ribosome. Cell 106, 233-241.](#)

[Zhang, Z., and Dietrich, F.S. \(2005\). Identification and characterization of upstream open reading frames \(uORF\) in the 5' untranslated regions \(UTR\) of genes in *Saccharomyces cerevisiae*. Curr Genet 48, 77-87.](#)

[Zhou, C., Arslan, F., Wee, S., Krishnan, S., Ivanov, A.R., Oliva, A., Leatherwood, J., and Wolf, D.A. \(2005\). PCI proteins eIF3e and eIF3m define distinct translation initiation factor 3 complexes. BMC Biol 3, 14.](#)

[Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. \(1995\). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. EMBO J 14, 4065-4072.](#)

[Zuberek, J., Wyslouch-Cieszynska, A., Niedzwiecka, A., Dadlez, M., Stepinski, J., Augustyniak, W., Gingras, A.C., Zhang, Z., Burley, S.K., Sonenberg, N., et al. \(2003\). Phosphorylation of eIF4E attenuates its interaction with mRNA 5' cap analogs by electrostatic repulsion: intein-mediated protein ligation strategy to obtain phosphorylated protein. RNA 9, 52-61.](#)