



**HAL**  
open science

## **Rôle du facteur d'initiation e/aIF2 dans le démarrage de la traduction chez les Eucaryotes et chez les archées.**

Laure Yatime

### ► **To cite this version:**

Laure Yatime. Rôle du facteur d'initiation e/aIF2 dans le démarrage de la traduction chez les Eucaryotes et chez les archées.. Génétique des populations [q-bio.PE]. Ecole Polytechnique X, 2005. Français. <NNT : >. <pastel-00001539>

**HAL Id: pastel-00001539**

**<https://pastel.hal.science/pastel-00001539v1>**

Submitted on 27 Jul 2010

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

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



HAL Authorization

THÈSE  
pour obtenir le titre de  
DOCTEUR de l'ÉCOLE POLYTECHNIQUE  
Discipline : Biologie

Présentée et soutenue publiquement par  
Laure YATIME  
Le 16 novembre 2005

**Rôle du facteur d'initiation e/aIF2 dans le démarrage de la traduction chez  
les Eucaryotes et chez les Archées**

Jury composé de

Pr. Sylvain BLANQUET	
Dr. Richard BUCKINGHAM	Rapporteur
Dr. Jacqueline CHERFILS	Rapporteur
Dr. Richard GIEGE	
Dr. Yves MECHULAM	Directeur de thèse
Dr. Emmanuelle SCHMITT	







## RÉSUMÉ

Le facteur hétérotrimérique e/aIF2 joue un rôle central dans le démarrage de la traduction chez les Eucaryotes et chez les Archées. Il conduit l'ARNt initiateur méthionylé jusqu'au ribosome et assure la spécificité de sélection du codon de démarrage sur l'ARNm. La structure cristallographique d'aIF2 $\gamma$  de l'archée *Pyrococcus abyssi*, précédemment résolue au laboratoire, a révélé une très forte homologie entre aIF2 $\gamma$ , qui constitue le cœur de l'hétérotrimère, et le facteur d'élongation bactérien EF1A. Cependant,  $\gamma$  possède des caractéristiques structurales propres qui pourraient être responsables de sa spécificité d'action dans le démarrage de la traduction.

Une étape cruciale de ce travail a consisté à développer un test de suivi *in vitro* de l'association d'aIF2 au Met-ARNt<sup>Met</sup>. Ce test a permis d'évaluer l'importance des caractéristiques du Met-ARNt<sup>Met</sup> et la contribution de chacune des sous-unités de l'hétérotrimère dans la formation du complexe aIF2:Met-ARNt<sup>Met</sup>. Ainsi, il a été montré que le résidu méthionine constitue un déterminant majeur dans la reconnaissance du Met-ARNt<sup>Met</sup> par aIF2. D'autre part, il apparaît que la sous-unité  $\gamma$  seule est effectivement capable de lier l'ARNt selon un mode similaire à celui observé pour EF1A mais avec une affinité considérablement réduite par rapport à celle de l'hétérotrimère. Nous avons montré que la présence de la sous-unité  $\alpha$  était nécessaire pour retrouver une affinité optimale vis-à-vis de l'ARNt tandis que la sous-unité  $\beta$  ne semble pas jouer de rôle dans cette liaison. L'utilisation d'une stratégie de découpage d' $\alpha$  en domaines séparés a montré que c'est le domaine 3 d'aIF2 $\alpha$  qui lie la sous-unité  $\gamma$  par l'intermédiaire d'une boucle du domaine 2 de  $\gamma$ . De plus, le dimère  $\alpha_{D3}\gamma$  semble nécessaire et suffisant pour retrouver une affinité pour l'ARNt comparable à celle du facteur natif.

Dans un second temps, des cristaux de la sous-unité  $\alpha$  entière et d'une forme tronquée correspondant aux domaines 2 et 3 ont été obtenus. Les structures d' $\alpha_{D2-3}$  et d' $\alpha$  complet ont été résolues à respectivement 2.26 et 3.37 Å de résolution. L'analyse du modèle structural a révélé une mobilité du domaine 3 d' $\alpha$  par rapport au bloc rigide formé par les domaines 1 et 2. La comparaison de séquences d'e/aIF2 $\alpha$  a montré que les zones de conservation d' $\alpha$  se situaient principalement dans le domaine 1 et dans le domaine 3 de la protéine, qui possèdent tous les deux des propriétés générales de liaison des ARNs. Le domaine 1 d' $\alpha$  pourrait ainsi interagir avec un autre partenaire de type ARN du démarrage de la traduction, tel que l'ARNm ou l'ARN ribosomal.

Finalement, des cristaux d'aIF2 $\alpha\gamma$  de *Sulfolobus solfataricus* ont été obtenus et la structure de l'hétérodimère a été résolue à 3.0 Å. Cette structure a confirmé les données biochimiques précédemment obtenues : le domaine 3 de la sous-unité  $\alpha$  interagit avec le domaine 2 de  $\gamma$ , au niveau de la boucle L1 précédemment caractérisée. L'analyse de cette structure a révélé pour la première fois une conformation des régions Switch de  $\gamma$  similaire à celle observée au sein du complexe EF1A:GDPNP:ARNt, ce qui permet d'expliquer la GTP-dépendance de la fixation du Met-ARNt<sup>Met</sup> par aIF2. La comparaison de cette structure à celle d'EF1A suggère que seule  $\gamma$  pourrait être en contact avec l'ARNt au sein de l'hétérodimère  $\alpha\gamma$ . Le renforcement de l'affinité pour l'ARNt observé en présence d' $\alpha$  nous a conduit à envisager un rôle possible d' $\alpha$  dans l'établissement des conformations observées pour les régions Switch dans la structure d'aIF2 $\gamma$ .

## SUMMARY

Eukaryotic and archaeal initiation factors 2 (e/aIF2) are heterotrimeric proteins ( $\alpha\beta\gamma$ ) that supply the small ribosomal subunit with methionylated initiator tRNA, therefore insuring specific recognition of the start codon on mRNA. The structure of aIF2 $\gamma$  from the archaeon *P. abyssi* was previously solved in our laboratory. The  $\gamma$  subunit, which forms the core of the heterotrimer, is a close structural homologue of bacterial elongation factor EF1A. However, it displays specific features that could account for its specific role in translation initiation.

One of the main goals of this work has been to design an *in vitro* assay to follow the association between aIF2 and Met-tRNA<sub>i</sub><sup>Met</sup>. This test enabled us to identify determinants of Met-tRNA<sub>i</sub><sup>Met</sup> important for its recognition by aIF2 and to investigate the role of each subunit of aIF2 in tRNA binding. On the one hand, the methionine moiety of initiator tRNA is a crucial determinant for Met-tRNA<sub>i</sub><sup>Met</sup> recognition by aIF2 while the other nucleotidic determinants have only a minor role. On the other hand, the  $\gamma$  subunit alone is able to bind Met-tRNA<sub>i</sub><sup>Met</sup> in a manner similar to that of EF1A. But its affinity for the tRNA molecule is strongly reduced in comparison to the affinity of the intact heterotrimer. Therefore, at least one other subunit must play a role in this interaction. Indeed, the  $\alpha$  subunit is required to have a full tRNA binding while  $\beta$  has no role in this interaction. Within the  $\alpha$  subunit, domain 3 binds the  $\gamma$  subunit *via* an idiosyncratic loop located in domain 2 of aIF2 $\gamma$ . Moreover, the  $\alpha_{D3}\gamma$  heterodimer is necessary and sufficient to have an optimal affinity for Met-tRNA<sub>i</sub><sup>Met</sup>.

In a second step, we obtained crystals of the intact aIF2 $\alpha$  and of a truncated version of  $\alpha$  formed of only domains 2 and 3. The structures of  $\alpha_{D2-3}$  and  $\alpha$  were solved at 2.26 and 3.37 Å resolution, respectively. The three-domain organisation of  $\alpha$  is conserved in Eukarya and Archaea. Domains 1 and 2 form a rigid body which is linked to a third mobile domain. Sequence comparisons established that the most conserved regions in aIF2 $\alpha$  lie at opposite sides of the protein, within domain 1 and domain 3. Both domains show general RNA binding properties. Thereby, domain 1 could interact with either rRNA or mRNA on the ribosome.

Finally, crystals of aIF2 $\alpha\gamma$  from *Sulfolobus solfataricus* were obtained and the structure of the heterodimer was solved and refined to 3.0 Å resolution. This structure confirmed biochemical data previously obtained:  $\alpha_{D3}$  connects domain 2 of the  $\gamma$  subunit *via* the L1 loop of  $\gamma$ . For the first time in an aIF2 $\gamma$  structure, conformation of the two switch regions of  $\gamma$  involved in GTP binding are similar to those encountered in the EF1A:GTP:Phe-tRNA<sup>Phe</sup> complex. Comparison with the EF1A structure suggests that only the  $\gamma$  subunit of the aIF2 $\alpha\gamma$  heterodimer contacts tRNA. Because the  $\alpha$  subunit markedly reinforces the affinity of the  $\gamma$  subunit for tRNA, a contribution of  $\alpha$  to the switch movements observed in the  $\gamma$  structure is considered.

# *REMERCIEMENTS*



Je tiens à remercier Sylvain Blanquet et Pierre Plateau pour m'avoir accueillie au sein du Laboratoire de Biochimie et offert l'opportunité de réaliser mon travail de thèse au sein d'une équipe particulièrement chaleureuse et dynamique.

Je remercie également Richard Buckingham et Jacqueline Cherfils pour avoir accepté d'être rapporteurs de cette thèse, ainsi que Richard Giegé pour avoir bien voulu examiner ce travail.

Je souhaite ensuite exprimer ma plus profonde reconnaissance à Yves Mechulam et Emmanuelle Schmitt qui m'ont encadrée et guidée tout au long de ce travail de thèse avec compétence et efficacité, sans jamais se départir de leur bonne humeur.

Mes remerciements vont également à Gil Craescu, de l'Institut Curie d'Orsay, pour son aide précieuse sur les expériences de dichroïsme circulaire, ainsi qu'à l'équipe du Laboratoire de Chimie des Mécanismes Réactionnels de l'Ecole Polytechnique, pour les analyses par spectrométrie de masse. Un grand merci également à Thomas Simonson pour la partie « linguistique ».

Je souhaite également remercier tous les membres du laboratoire que j'ai pu côtoyer au cours de ces trois années : mes compagnons de paillasse, les « B2 », Christine Schurdevin, Michel Panvert et Laurent Guillon (encore un chimiste !), ma voisine de bureau Laura Fioni mon ex-co-bureau Thanh Ha-Duong (sans qui les pauses thé-chocolat n'auraient pas été les mêmes !), Myriam Lazard, Michel Fromant (mingalar pa ?), Marc Dauplais, Alfonso Jaramillo, Catherine Teillet et Françoise Humbert pour avoir supporté mes visites incessantes ces derniers mois, Annick Carpentier, Pascal Delplanche, Guillaume Heugas, Catherine Morais, Romary Perrin, Agathe Tarze, David Mignon, Eric Madigou, François Cluzeau, Codjo Hountondji, Florence Delort.

Je profite également de cette occasion pour remercier tous les thésards et post-doc du laboratoire pour les moments passés ensemble (et nos fameux repas jeunes !) : Ioana Grigoras, Laurent Giorgi, les bioinformaticiens Josselin Noirel, Guillaume Launay, Damien Thompson, Pablo Tortosa, Marcel Schmidt, Anne Lopes et Alexey Aleksandrov, sans oublier les « anciens » : Thibaut Crépin, Lionel Vial, Olivier Gueldry et Olga Soutourina.

Enfin, je remercie ma famille et mes proches, et en particulier ma mère qui m'a soutenue avec patience et dévouement durant ces années de thèse.



*A mon père*



# *SOMMAIRE*



	PAGES
<i>LISTE DES FIGURES ET TABLEAUX</i>	1
<i>ABRÉVIATIONS</i>	9
<i>Chapitre I : INTRODUCTION</i>	11
<i>Le démarrage de la traduction. Similitudes et divergences au sein du monde vivant</i>	
<i>Partie A : Le démarrage de la traduction bactérienne</i>	13
1) <i>Les étapes du mécanisme</i>	13
2) <i>Formation du complexe de prédémarrage</i>	17
1. <i>Activités anti-associatrices d'IF3 et d'IF1</i>	17
2. <i>Liaison de l'ARNm</i>	21
3. <i>Liaison de l'ARNt</i>	22
3) <i>Interaction codon-anticodon et décodage</i>	29
4) <i>Association des sous-unités ribosomales</i>	31
<i>Partie B : Le démarrage de la traduction eucaryotique</i>	36
1) <i>Les étapes du mécanisme</i>	36
2) <i>Formation du complexe de prédémarrage 43S</i>	40
1. <i>Les partenaires protéiques du complexe de prédémarrage</i>	40
2. <i>Activité anti-associatrice d'eIF3, eIF1 et eIF1A</i>	43
3. <i>eIF6 : un cas ambigu ?</i>	46
3) <i>Recrutement de l'ARNm et balayage de sa séquence</i>	47
1. <i>Reconnaissance de la coiffe en 5'</i>	48
2. <i>Reconnaissance de la queue poly-(A) en 3'</i>	52

4) Reconnaissance du codon AUG et hydrolyse du GTP	55
5) Association des sous-unités ribosomales	57
6) Autres modes de démarrage	61
<b>Partie C : Le démarrage de la traduction archéenne</b>	63
1) La traduction chez les Archées : un mécanisme hybride	63
2) Démarrage de la traduction archéenne : similitudes et divergences avec les autres domaines	64
1. L'ARNm	65
2. Les facteurs de démarrage	65
3) Conclusion	67
<b>Chapitre II : LE FACTEUR D'INITIATION e/aIF2</b>	69
<b>Fonctionnalités et Organisation Structurale</b>	
<b>Partie A : Rôle général du facteur de démarrage e/aIF2</b>	69
1) Description fonctionnelle	69
2) Les protéines G : un mécanisme universel	69
1. Mécanisme du changement conformationnel	71
2. Hydrolyse du GTP, GAP-assistée	72
3. Recyclage GDP → GTP, GEF-assisté	73
<b>Partie B : La sous-unité <math>\gamma</math> du facteur e/aIF2</b>	74
1) Structure de la sous-unité $\gamma$	74
2) Zones spécifiques au facteur de démarrage	76
3) Modèle de positionnement de l'ARNt	77
4) Mouvement des régions Switch : comparaison d'e/aIF2 $\gamma$ avec EF1A	79
5) Perspectives	81

<i>Partie C : La sous-unité <math>\beta</math> du facteur e/aIF2</i>	82
1) <i>Généralités</i>	82
2) <i>Structure de la sous-unité <math>\beta</math></i>	83
3) <i>Fonction d'e/aIF2<math>\beta</math></i>	84
4) <i>Modèle de positionnement d'aIF2<math>\beta</math> sur aIF2<math>\gamma</math></i>	86
5) <i>Le domaine N-terminal d'eIF2<math>\beta</math>: une spécificité eucaryotique</i>	87
6) <i>eIF5 : une GAP classique ?</i>	88
<i>Partie D : La sous-unité <math>\alpha</math> du facteur e/aIF2</i>	91
1) <i>Généralités</i>	91
2) <i>Structure de la sous-unité <math>\alpha</math></i>	92
3) <i>Le recyclage GDP <math>\rightarrow</math> GTP chez les Eucaryotes</i>	93
<b>OBJECTIFS</b>	95
<b>Chapitre III : CARTOGRAPHIE FONCTIONNELLE DU FACTEUR aIF2</b>	97
<i>Partie A : Reconnaissance spécifique de l'ARNt initiateur par le facteur aIF2</i>	97
1) <i>Clonage des gènes codant pour les 3 sous-unités du facteur aIF2 de P. abyssii</i>	97
2) <i>Purification des 3 sous-unités du facteur aIF2 de P. abyssii</i>	98
1. <i>Mode de surproduction adopté</i>	98
2. <i>Purification de Pa-aIF2<math>\alpha</math>, Pa-aIF2<math>\beta</math> et Pa-aIF2<math>\gamma</math></i>	99

<b>4) Mesure de l'affinité de Pa-aIF2 pour l'ARNt</b>	100
1. Principe de la méthode	100
2. Mise en œuvre expérimentale	103
3. Résultats préliminaires et mesure du Kd pour le complexe Pa-aIF2:Met-ARNt <sub>i</sub> <sup>Met</sup>	103
4. Spécificité de Pa-aIF2 pour l'ARNt initiateur	106
<b>5) Rôle des différentes sous-unités de Pa-aIF2 dans la liaison de l'ARNt</b>	110
1. Rôle de Pa-aIF2 $\gamma$	110
2. Rôle de Pa-aIF2 $\alpha$	111
<b>Partie B : Cartographie des zones d'interactions entre les sous-unités du facteur Pa-aIF2</b>	115
<i>La sous-unité <math>\gamma</math>, maillon central de l'assemblage</i>	
<b>1) Obtention des mutants de la sous-unité <math>\gamma</math></b>	115
1. Obtention des différents plasmides mutés	115
2. Production et purification des mutants de Pa-aIF2 $\gamma$	117
3. Production et purification des mutants du domaine du zinc	117
<b>2) Tests d'assemblages sur gels natifs</b>	119
<b>3) Vérification de la stabilité structurale du mutant <math>\gamma_{\Delta L1}</math></b>	124
<b>4) Validation du modèle de liaison de l'ARNt fourni par EF1A</b>	127
<b>Partie C : Cartographie des zones d'interactions entre les sous-unités du facteur Pa-aIF2</b>	131
<i>La sous-unité <math>\alpha</math>, un partenaire essentiel</i>	
<b>1) Production des domaines séparés de Pa-aIF2<math>\alpha</math></b>	131
<b>2) Tests d'assemblages avec la sous-unité <math>\gamma</math></b>	133
<b>3) Implications fonctionnelles</b>	135
<b>Conclusions au Chapitre III</b>	137

<b>Chapitre IV : ÉTUDE STRUCTURALE DE LA SOUS-UNITÉ <math>\alpha</math>IF2<math>\alpha</math></b>	<b>138</b>
<b>Partie A : Résolution de la structure de la sous-unité <math>\alpha</math>IF2<math>\alpha</math> de l'archée <i>Pyrococcus abyssi</i></b>	<b>138</b>
<b>1) Principe de la cristallographie par diffraction des rayons X</b>	<b>138</b>
<b>2) Obtention des cristaux d'<math>\alpha</math>IF2<math>\alpha_{D2-3}</math> de <i>P. abyssi</i></b>	<b>139</b>
<b>3) Acquisition et traitement des données</b>	<b>142</b>
1. Source de rayons X	142
2. Préparation des échantillons	143
3. Acquisition et traitement des données pour la protéine native	143
<b>4) Résolution d'une structure par la diffusion anormale multi-longueur d'onde (MAD)</b>	<b>145</b>
1. Principe	145
2. Obtention de la protéine $\alpha_{D2-3}$ sélénométhionylée, cristallisation et enregistrement des données	147
3. Résolution de la structure	149
<b>5) Affinement du modèle</b>	<b>150</b>
<b>6) Résolution de la structure d'<math>\alpha</math>IF2<math>\alpha</math> de <i>P. abyssi</i></b>	<b>153</b>
1. Cristallisation d' $\alpha$ IF2 $\alpha$ et acquisition des données	153
2. Principe de la méthode du remplacement moléculaire	154
3. Résolution de la structure	156
<b>Partie B : Structure de la sous-unité <math>\alpha</math>IF2<math>\alpha</math> de l'archée <i>Pyrococcus abyssi</i></b>	<b>160</b>
<b>1) Description générale</b>	<b>160</b>
1. Structure de Pa- $\alpha$ IF2 $\alpha_{D2-3}$	160
2. Structure de la protéine intacte Pa- $\alpha$ IF2 $\alpha$	161
3. Orientation relative des 3 domaines de la sous-unité $\alpha$	162

2) <i>Homologies structurales et de séquence</i>	166
1. Homologies structurales	166
2. Homologies de séquence	167
3. Implications fonctionnelles	168
<i>Chapitre V : ÉTUDE FONCTIONNELLE ET STRUCTURALE DE L'HÉTÉRODIMÈRE <math>\alpha IF2\alpha\gamma</math></i>	172
<i>Partie A : L'hétérodimère <math>\alpha IF2\alpha_{D3}\gamma</math> de <i>P. abyss</i></i>	172
1) <i>Cristallisation d'<math>\alpha IF2\alpha_{D3}\gamma</math> de <i>P. abyss</i></i>	172
2) <i>Enregistrement et traitement des données</i>	174
3) <i>Tentatives pour résoudre la structure</i>	175
<i>Partie B : Caractérisation biochimique du facteur <math>\alpha IF2</math> de <i>Sulfolobus solfataricus</i></i>	178
1) <i>Introduction</i>	178
2) <i>Clonage des gènes codant pour les 3 sous-unités d'<math>\alpha IF2</math> de <i>S. solfataricus</i> et purification des protéines</i>	178
3) <i>Affinité vis-à-vis du <math>Met-ARNt_i^{Met}</math> et rôle des différentes sous-unités</i>	180
<i>Partie C : Résolution de la structure de l'hétérodimère <math>\alpha IF2\alpha\gamma</math> de l'archée <i>Sulfolobus solfataricus</i></i>	183
1) <i>Cristallisation de <i>Ss</i>-<math>\alpha IF2\alpha\gamma</math> et enregistrement des données</i>	183
2) <i>Résolution de la structure et affinement du modèle</i>	185

<i>Partie D : Structure de l'hétérodimère <math>\alpha</math>IF2<math>\alpha\gamma</math> de l'archée <i>Sulfolobus solfataricus</i></i>	189
1) <i>Description générale</i>	189
1. Organisation structurale de l'hétérodimère	189
2. Zone d'interaction entre les deux sous-unités	190
3. Poche de fixation des nucléotides guanine	192
4. Les régions Switch I et Switch II	194
2) <i>Positionnement de l'ARNt au sein du complexe Ss-<math>\alpha</math>IF2<math>\alpha\gamma</math></i>	196
1. Modèle de positionnement	196
2. Rôle de la sous-unité $\alpha$ dans la liaison de l'ARNt	198
3) <i>Les sous-unités <math>\alpha</math> et <math>\gamma</math> : un rôle concerté au sein de l'hétérotrimère ?</i>	199
 <i>CONCLUSION</i>	 203
 <i>MATÉRIEL ET MÉTHODES</i>	 208
 <i>ANNEXES</i>	 237
 <i>RÉFÉRENCES BIBLIOGRAPHIQUES</i>	 241
 <i>LISTE DES TRAVAUX</i>	 263



*LISTE DES TABLEAUX  
ET FIGURES*



	PAGES
<b>INTRODUCTION</b>	
<i>Figure 1.1 :</i>	<i>Vue générale des sous-unités ribosomales</i> 14
<i>Figure 1.2 :</i>	<i>Mécanisme général du démarrage de la traduction bactérienne</i> 17
<i>Figure 1.3 :</i>	<i>Structure tridimensionnelle des domaines d'IF3</i> 18
<i>Figure 1.4 :</i>	<i>Modèle de positionnement d'IF3 sur la sous-unité 30S</i> 19
<i>Figure 1.5 :</i>	<i>Structure tridimensionnelle d'IF1 sur la sous-unité 30S</i> 20
<i>Figure 1.6 :</i>	<i>Cheminement de l'ARN<sup>m</sup> au sein de la sous-unité 30S</i> 22
<i>Figure 1.7 :</i>	<i>Caractéristiques distinctives des ARN<sup>t</sup><sup>Met</sup> initiateur et élongateur</i> 24
<i>Figure 1.8 :</i>	<i>Organisation structurale d'IF2</i> 25
<i>Figure 1.9 :</i>	<i>Zones d'interaction d'IF2 sur le Met-ARN<sup>t</sup><sub>f</sub><sup>Met</sup></i> 26
<i>Figure 1.10 :</i>	<i>Position d'IF2 au sein du complexe 70S par cryo-EM</i> 27
<i>Figure 1.11 :</i>	<i>Recouvrement du site ribosomal A par IF1 sur la sous-unité 30S</i> 29
<i>Figure 1.12 :</i>	<i>Interactions entre IF3 et l'ARN<sup>t</sup> initiateur au site P</i> 31
<i>Figure 1.13 :</i>	<i>Positionnement d'IF2 à l'interface 30S – 50S au sein du complexe d'initiation 70S par cryo-EM</i> 33
<i>Figure 1.14 :</i>	<i>Positionnement d'IF2 sur chacune des sous-unités 30S et 50S au sein du complexe d'initiation 70S par cryo-EM</i> 34
<i>Figure 1.15 :</i>	<i>Mécanisme général du démarrage de la traduction eucaryotique</i> 39
<i>Figure 1.16 :</i>	<i>Modèle tridimensionnel du MFC</i> 41
<i>Figure 1.17 :</i>	<i>Structure tridimensionnelle d'eIF1</i> 42
<i>Figure 1.18 :</i>	<i>Structures tridimensionnelles d'e/aIF1A et d'IF1</i> 42

Figure 1.19 :	Modèle de positionnement d'eIF3 sur la sous-unité 40S	45
Figure 1.20 :	Structure tridimensionnelle du complexe eIF4E-coiffe de l'ARNm	49
Figure 1.21 :	Cartographie des interactions d'eIF4G	50
Figure 1.22 :	Structure tridimensionnelle d'eIF4A	51
Figure 1.23 :	Structure tridimensionnelle de 2 des domaines RRM de PABP	53
Figure 1.24 :	Modèle de circularisation de l'ARNm	54
Figure 1.25 :	Mouvements conformationnels au sein d'e/aIF5B lors de l'hydrolyse du GTP	59
Tableau 1.26 :	Facteurs de démarrage de la traduction dans les trois domaines du monde vivant	64
Figure 1.27 :	Modèle de démarrage de la traduction chez LUCA	67
 <b>CHAPITRE II</b> 		
Figure 2.1 :	Structure du domaine G commun à toutes les GTPases	70
Figure 2.2 :	Mouvement conformationnel des régions Switch du domaine G des GTPases	71
Figure 2.3 :	Changements conformationnels au sein d'EF1A après hydrolyse du GTP	72
Tableau 2.4 :	Comparaison structurale d'EF1A et d'e/aIF2γ : rmsd déduits de la comparaison de leurs domaines séparés	75
Figure 2.5 :	Structure tridimensionnelle d'aIF2γ de <i>P. abyssi</i>	75
Figure 2.6 :	Tétracoordination de l'ion Zn <sup>2+</sup> au sein du domaine Zinc d'aIF2γ de <i>P. abyssi</i>	77
Figure 2.7 :	Modèle de positionnement du Phe-ARNt <sup>Phe</sup> sur la structure d'aIF2γ de <i>P. abyssi</i>	78

<i>Figure 2.8 :</i>	<i>Mouvements conformationnels au sein d'aIF2<math>\gamma</math> de <i>P. abyssi</i> entre les formes GDP-Mg<sup>2+</sup> et GDPNP-Mg<sup>2+</sup></i>	80
<i>Figure 2.9 :</i>	<i>Organisation structurale d'e/aIF2<math>\beta</math></i>	82
<i>Figure 2.10 :</i>	<i>Structure tridimensionnelle d'aIF2<math>\beta</math> de <i>M. thermoautotrophicum</i></i>	83
<i>Figure 2.11 :</i>	<i>Modèle de positionnement d'aIF2<math>\beta</math> sur la sous-unité <math>\gamma</math> du facteur de démarrage</i>	86
<i>Figure 2.12 :</i>	<i>Cartographie des interactions d'eIF2<math>\beta</math> avec eIF5 et eIF2B</i>	88
<i>Figure 2.13 :</i>	<i>Organisation structurale d'e/aIF2<math>\alpha</math></i>	91
<i>Figure 2.14 :</i>	<i>Structure tridimensionnelle des deux domaines N-terminaux d'eIF2<math>\alpha</math> de levure</i>	92

### CHAPITRE III

<i>Tableau 3.1 :</i>	<i>Paramètres physico-chimiques des 3 sous-unités du facteur aIF2 de <i>P. abyssi</i></i>	97
<i>Figure 3.2 :</i>	<i>Caractérisation des 3 sous-unités du facteur aIF2 de <i>P. abyssi</i> par migration électrophorétique</i>	100
<i>Figure 3.3 :</i>	<i>Courbes de déacylation du Met-ARNt<sub>f</sub><sup>Met</sup> en présence du facteur aIF2 de <i>P. abyssi</i></i>	104
<i>Figure 3.4 :</i>	<i>Courbe de mesure de la constante de dissociation pour le complexe Met-ARNt<sub>f</sub><sup>Met</sup> - aIF2 de <i>P. abyssi</i></i>	104
<i>Figure 3.5 :</i>	<i>Comparaison des ARNts initiateurs d'<i>E. coli</i> et de <i>P. abyssi</i></i>	106
<i>Figure 3.6 :</i>	<i>Schéma des différents ARNts, natifs et mutants, utilisés dans le test de protection contre la déacylation</i>	108
<i>Tableau 3.7 :</i>	<i>Identification des déterminants de l'ARNt initiateur méthionylé reconnu par le facteur aIF2 de <i>P. abyssi</i></i>	109

<i>Tableau 3.8 :</i>	<i>Rôle de la sous-unité <math>\gamma</math> dans la liaison du Met-ARNt<sup>Met</sup></i>	<i>110</i>
<i>Figure 3.9 :</i>	<i>Vérification de la stoechiométrie des complexes utilisés dans le test de protection par électrophorèse native</i>	<i>111</i>
<i>Tableau 3.10 :</i>	<i>Rôle des sous-unités <math>\alpha</math> et <math>\beta</math> dans la liaison de l'ARNt initiateur méthionylé</i>	<i>112</i>
<i>Figure 3.11 :</i>	<i>Retard de migration de l'ARNt<sup>Met</sup> sur gel d'électrophorèse en présence de la sous-unité <math>\alpha</math> du facteur aIF2</i>	<i>112</i>
<i>Figure 3.12 :</i>	<i>Propriétés générales de liaison des acides nucléiques par aIF2<math>\alpha</math></i>	<i>113</i>
<i>Figure 3.13 :</i>	<i>Cartographie des régions d'aIF2<math>\gamma</math> de <i>P. abyssi</i> soumises à une analyse mutationnelle</i>	<i>116</i>
<i>Figure 3.14 :</i>	<i>Purification des mutants d'aIF2<math>\gamma</math> de <i>P. abyssi</i></i>	<i>117</i>
<i>Figure 3.15 :</i>	<i>Purification du mutant aIF2<math>\gamma_{\Delta Z_n}</math> sous forme hétérodimérique <math>\alpha\gamma</math></i>	<i>119</i>
<i>Figure 3.16 :</i>	<i>Assemblages par électrophorèse sur gel natif des mutants de boucles d'aIF2<math>\gamma</math> avec les sous-unités <math>\alpha</math> et <math>\beta</math></i>	<i>120</i>
<i>Figure 3.17 :</i>	<i>Assemblages par électrophorèse sur gel natif des mutants ponctuels d'aIF2<math>\gamma</math> avec les sous-unités <math>\alpha</math> et <math>\beta</math></i>	<i>121</i>
<i>Figure 3.18 :</i>	<i>Association de la sous-unité <math>\beta</math> avec le mutant aIF2<math>\gamma_{\Delta Z_n}</math> lors de la purification du complexe <math>\alpha\gamma_{\Delta Z_n}</math> sur TALON</i>	<i>122</i>
<i>Tableau 3.19 :</i>	<i>Capacité de liaison à <math>\alpha</math> et à <math>\beta</math> des différents mutants de la sous-unité <math>\gamma</math> du facteur aIF2 de <i>P. abyssi</i></i>	<i>123</i>
<i>Figure 3.20 :</i>	<i>Spectres de dichroïsme circulaire d'aIF2<math>\gamma</math> natif et du mutant <math>\gamma_{\Delta L1}</math></i>	<i>125</i>
<i>Figure 3.21 :</i>	<i>Dénaturation thermique d'aIF2<math>\gamma</math> natif et du mutant <math>\gamma_{\Delta L1}</math></i>	<i>127</i>
<i>Tableau 3.22 :</i>	<i>Propriétés de liaison de l'ARNt initiateur méthionylé par les mutants d'aIF2<math>\gamma</math> de <i>P. abyssi</i></i>	<i>128</i>

<i>Figure 3.23 :</i>	<i>Modèle de positionnement de la tige acceptrice de l'ARNt au sein du domaine 2 d'aIF2<math>\gamma</math></i>	129
<i>Figure 3.24 :</i>	<i>Production des domaines séparés d'aIF2<math>\alpha</math> de <i>P. abyssi</i></i>	131
<i>Figure 3.25 :</i>	<i>Purification des domaines séparés d'aIF2<math>\alpha</math> de <i>P. abyssi</i></i>	132
<i>Figure 3.26 :</i>	<i>Assemblages des domaines séparés d'aIF2<math>\alpha</math> avec la sous-unité <math>\gamma</math> par électrophorèse sur gel natif</i>	133
<i>Figure 3.27 :</i>	<i>Assemblages des domaines séparés d'aIF2<math>\alpha</math> avec la sous-unité <math>\gamma</math> par chromatographie d'exclusion</i>	134
<i>Tableau 3.28 :</i>	<i>Capacité de liaison d'aIF2<math>\gamma</math> par les domaines séparés de la sous-unité <math>\alpha</math></i>	135
<i>Tableau 3.29 :</i>	<i>Rôle des différents domaines d'aIF2<math>\alpha</math> dans le renforcement de l'affinité d'aIF2<math>\gamma</math> pour l'ARNt initiateur méthionylé</i>	135
<i>Figure 3.30 :</i>	<i>Propriétés générales de liaison des acides nucléiques par les domaines 1 et 3 d'aIF2<math>\alpha</math></i>	136

## CHAPITRE IV

<i>Figure 4.1 :</i>	<i>Applications de la technique de la diffusion de vapeur</i>	140
<i>Tableau 4.2 :</i>	<i>Conditions de cristallisation d'aIF2<math>\alpha_{D2-3}</math> de <i>P. abyssi</i></i>	141
<i>Figure 4.3 :</i>	<i>Cristaux de la protéine aIF2<math>\alpha_{D2-3}</math> de <i>P. abyssi</i></i>	141
<i>Figure 4.4 :</i>	<i>Purification de la forme aIF2<math>\alpha_{D2-3}</math> sélénométhionylée</i>	148
<i>Tableau 4.5 :</i>	<i>Conditions d'enregistrement des données de diffraction relatives à aIF2<math>\alpha_{D2-3}</math> de <i>P. abyssi</i></i>	149
<i>Tableau 4.6 :</i>	<i>Statistiques du traitement des données de diffraction relatives à aIF2<math>\alpha_{D2-3}</math> de <i>P. abyssi</i></i>	150
<i>Tableau 4.7 :</i>	<i>Données cristallographiques du modèle structural d'aIF2<math>\alpha_{D2-3}</math> de <i>P. abyssi</i></i>	151

<i>Figure 4.8 :</i>	<i>Diagramme de Ramachandran du modèle d'<math>aIF2\alpha_{D2-3}</math> de <i>P. abyss</i> affiné à 2.26 Å de résolution</i>	152
<i>Tableau 4.9 :</i>	<i>Conditions d'enregistrement des données de diffraction relatives à <math>aIF2\alpha</math> de <i>P. abyss</i></i>	154
<i>Tableau 4.10 :</i>	<i>Statistiques du traitement des données de diffraction relatives à <math>aIF2\alpha</math> de <i>P. abyss</i></i>	154
<i>Tableau 4.11 :</i>	<i>Données cristallographiques du modèle structural d'<math>aIF2\alpha</math> de <i>P. abyss</i></i>	157
<i>Figure 4.12 :</i>	<i>Diagramme de Ramachandran du modèle d'<math>aIF2\alpha</math> de <i>P. abyss</i> affiné à 3.37 Å de résolution</i>	158
<i>Figure 4.13 :</i>	<i>Structure d'<math>aIF2\alpha_{D2-3}</math> de <i>P. abyss</i> à 2.26 Å de résolution</i>	160
<i>Figure 4.14 :</i>	<i>Structure d'<math>aIF2\alpha</math> de <i>P. abyss</i> à 3.37 Å de résolution</i>	162
<i>Figure 4.15 :</i>	<i>Orientation relative des 3 domaines d'<math>e/aIF2\alpha</math></i>	163
<i>Figure 4.16 :</i>	<i>Cartographie des interactions entre les domaines 1 et 2 d'<math>e/aIF2\alpha</math></i>	164
<i>Figure 4.17 :</i>	<i>Mobilité du domaine 3 d'<math>aIF2\alpha</math> par rapport aux domaines 1 et 2</i>	165
<i>Figure 4.18 :</i>	<i>Position des résidus conservés d'<math>e/aIF2\alpha</math></i>	168
<i>Figure 4.19 :</i>	<i>Alignement de séquences des <math>e/aIF2\alpha</math></i>	171

## CHAPITRE V

<i>Tableau 5.1 :</i>	<i>Conditions de cristallisation de l'hétérodimère <math>aIF2\alpha_{D3\gamma}</math> de <i>P. abyss</i></i>	172
<i>Figure 5.2 :</i>	<i>Analyse du contenu protéique des cristaux</i>	173
<i>Figure 5.3 :</i>	<i>Cristaux de l'hétérodimère <math>aIF2\alpha_{D3\gamma}</math> de <i>P. abyss</i></i>	174
<i>Tableau 5.4 :</i>	<i>Conditions d'enregistrement des données de diffraction relatives à l'hétérodimère <math>aIF2\alpha_{D3\gamma}</math> de <i>P. abyss</i></i>	175

Tableau 5.5 :	<i>Statistiques du traitement des données de diffraction relatives à l'hétérodimère <math>\alpha\text{IF}2\alpha_{\text{D}3}\gamma</math> de <i>P. abyssi</i></i>	175
Figure 5.6 :	<i>Modèle rigide non affiné de l'hétérodimère <math>\alpha\text{IF}2\alpha_{\text{D}3}\gamma</math> de <i>P. abyssi</i></i>	177
Tableau 5.7 :	<i>Paramètres physico-chimiques des 3 sous-unités du facteur <math>\alpha\text{IF}2</math> de <i>S. solfataricus</i></i>	178
Figure 5.8 :	<i>Etapes de purification du facteur <math>\alpha\text{IF}2</math> de <i>S. solfataricus</i></i>	179
Figure 5.9 :	<i>Migration électrophorétique de l'hétérotrimère et des deux hétérodimères natifs du facteur <math>\alpha\text{IF}2</math> de <i>S. solfataricus</i></i>	180
Figure 5.10 :	<i>Courbes de mesure de la constante de dissociation pour le complexe <math>\text{Met-ARNt}_f^{\text{Met}} - \alpha\text{IF}2</math> de <i>S. solfataricus</i></i>	181
Tableau 5.11 :	<i>Rôle des différentes sous-unités d'<math>\alpha\text{IF}2</math> de <i>S. solfataricus</i> dans la liaison de l'ARNt initiateur méthionylé</i>	182
Tableau 5.12 :	<i>Conditions de cristallisation de l'hétérodimère <math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i></i>	183
Figure 5.13 :	<i>Analyse du contenu protéique des cristaux</i>	183
Figure 5.14 :	<i>Cristaux de l'hétérodimère <math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i></i>	184
Tableau 5.15 :	<i>Conditions d'enregistrement des données de diffraction relatives à l'hétérodimère <math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i></i>	184
Tableau 5.16 :	<i>Statistiques du traitement des données de diffraction relatives à l'hétérodimère <math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i></i>	185
Figure 5.17 :	<i>Amélioration de la densité électronique autour de la molécule de <math>\text{GDPNP}</math> au cours de l'affinement</i>	186
Tableau 5.18 :	<i>Données cristallographiques du modèle structural d'<math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i></i>	187
Figure 5.19 :	<i>Diagramme de Ramachandran du modèle d'<math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i> affiné à 3.0 Å de résolution</i>	188
Figure 5.20 :	<i>Structure de l'hétérodimère <math>\alpha\text{IF}2\alpha\gamma</math> de <i>S. solfataricus</i> à</i>	189

---

	<b>3.0 Å de résolution</b>	
<i>Figure 5.21 :</i>	<i>Zone d'interaction entre les sous-unités <math>\alpha</math> et <math>\gamma</math> du facteur <math>\alpha</math>IF2 de <i>S. solfataricus</i></i>	<b>191</b>
<i>Figure 5.22 :</i>	<i>Poche de fixation des nucléotides guanine</i>	<b>193</b>
<i>Figure 5.23 :</i>	<i>Comparaison de la position des régions Switch du domaine G de <i>Ss</i>-<math>\alpha</math>IF2<math>\gamma</math>, de <i>Pa</i>-<math>\alpha</math>IF2<math>\gamma</math> et d'<i>EF1A</i></i>	<b>195</b>
<i>Figure 5.24 :</i>	<i>Modèle de positionnement du <i>Phe</i>-ARNt<sup><i>Phe</i></sup> sur la structure d'<math>\alpha</math>IF2<math>\alpha\gamma</math> de <i>S. solfataricus</i></i>	<b>197</b>
<i>Tableau 5.25 :</i>	<i>Affinité des hétérodimères <math>\alpha_{\text{mutés}}\gamma</math> pour l'ARNt initiateur méthionylé</i>	<b>199</b>
<i>Tableau 5.26 :</i>	<i>GTP-dépendance de la liaison de l'ARNt initiateur méthionylé par l'hétérodimère <math>\alpha</math>IF2<math>\alpha\gamma</math> de <i>S. solfataricus</i></i>	<b>201</b>

## CONCLUSION

<i>Figure 5.27 :</i>	<i>Modèle de positionnement de l'hétérodimère <math>\alpha</math>IF2<math>\alpha\gamma</math> de <i>S. solfataricus</i> sur la sous-unité ribosomale 30S en présence d'un ARNt initiateur au site P</i>	<b>207</b>
----------------------	---	------------

## MATÉRIEL ET MÉTHODES

<i>Figure 6.1 :</i>	<i>Vecteurs <i>pET3<math>\alpha</math>lp</i> et <i>pET15b<math>\alpha</math>lp</i></i>	<b>212</b>
<i>Tableau 6.2 :</i>	<i>Amorces pour la synthèse des mutants d'<math>\alpha</math>IF2<math>\gamma</math> (1)</i>	<b>217</b>
<i>Figure 6.3 :</i>	<i>Mutagène dirigée par PCR de chevauchement</i>	<b>218</b>
<i>Tableau 6.4 :</i>	<i>Amorces pour la synthèse des mutants d'<math>\alpha</math>IF2<math>\gamma</math> (2)</i>	<b>219</b>
<i>Figure 6.5 :</i>	<i>Mutagène dirigée par la technique Stratagene</i>	<b>220</b>
<i>Figure 6.6 :</i>	<i>Principe de la purification sur colonne d'affinité TALON</i>	<b>221</b>
<i>Figure 6.7 :</i>	<i>Courbe d'étalonnage de la colonne TSK DNA</i>	<b>224</b>
<i>Figure 6.8 :</i>	<i>Amorces pour la synthèse des domaines d'<math>\alpha</math>IF2<math>\alpha</math></i>	<b>226</b>

# *ABBREVIATIONS*



aaRS	aminoacyl-ARNt synthétase
ADN	acide désoxyribonucléique
aIF	facteur de démarrage de la traduction archaebactérienne
Amp	ampicilline
ARN	acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomal
ARNt	ARN de transfert
ATP/ADP	adénosine tri/diphosphate
BET	bromure d'éthidium
Cm	chloramphénicol
Cryo-EM	cryo-microscopie électronique
DO	densité optique
DTT	dithiothréitol
EDTA	acide éthylène-diamine tétra-acétique
EF1A	facteur d'élongation bactérien 1A (ou EF-Tu)
eIF	facteur de démarrage de la traduction eucaryotique
ESRF	European Synchrotron Radiation Facility
fMet-ARNt <sub>f</sub> <sup>Met</sup>	ARNt initiateur bactérien
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 de démarrage de la traduction bactérienne 1, 2 et 3
IPTG	isopropyl-β-D-thiogalactoside
IRES	Internal Ribosomal Entry Site
Kan	kanamycine
LUCA	Last Universal Common Ancestor
Met-ARNt <sub>i</sub> <sup>Met</sup>	ARNt initiateur eucaryotique ou archaebactérien

MetRS	methionyl-ARNt synthetase
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	polyéthylène glycol
Pi	Phosphate inorganique
PMSF	phenylmethylsulfonyl fluoride
PTC	Peptidyl-Transferase Center
RMN	Résonance Magnétique Nucléaire
rmsd	root mean square deviation (écart quadratique moyen)
SD	séquence de Shine-Dalagarno
SDS	sodium dodécyl sulfate
SDS-PAGE	electrophorèse sur gel de polyacrylamide en présence de SDS
Spec	spectinomycine
TBE	Tris-Borate-EDTA
TEMED	N,N,N,N'-tetra-methylethylenediamine
Tris	Tris(hydroxyméthyl)aminométhane
5'-UTR	Région 5' non traduite (UnTranlated Region)

# *Chapitre I : Introduction*

*Le démarrage de la traduction.  
Similitudes et divergences au sein  
du monde vivant.*



La traduction est un processus cellulaire fondamental au sein de tout organisme vivant. Elle correspond à la dernière étape de décodage de l'information génétique, portée par l'acide désoxyribonucléique (ADN). Ce processus de décodage se déroule en deux temps.

La première phase, appelée transcription, permet de segmenter le support de l'information génétique. En effet, chacune des longues molécules d'ADN double brin enroulées en hélice est transposée en divers fragments d'acide ribonucléique messager (ARNm). Chaque fragment mono brin ne comporte ainsi qu'un nombre limité de gènes, ou même un unique gène comme c'est le cas pour les ARNm des Eucaryotes. Chez ces derniers, l'ARNm primaire nouvellement synthétisé subit une étape supplémentaire dite d'épissage qui permet d'exciser certaines régions non codantes (ou *introns*). Le transcrit épissé est alors doté d'une coiffe en 5' et d'une extrémité 3' polyadénine pour former l'ARNm mature qui sera traduit.

La seconde phase, appelée traduction ou plus communément synthèse des protéines, consiste à passer de cet alphabet à 4 lettres (A, U, G, C) commun aux acides nucléiques, à un nouvel alphabet, composé de 20 lettres - les 20 acides aminés qui entrent dans la composition des protéines. Le code génétique constitue la clé qui permet de passer d'un langage à l'autre. Cette clé est matérialisée au sein des cellules par des molécules particulières : les ARN de transfert (ARNt). Ces petites molécules polynucléotidiques comportent une extrémité anticodon qui va s'associer avec les codons présents sur l'ARNm par complémentarité. A l'autre extrémité de la molécule, un acide aminé, qui correspond à la séquence de l'anticodon, a été fixé par une aminoacyl-ARNt synthétase. Il sera attaché au polypeptide naissant après appariement correct entre le codon et l'anticodon. Ainsi, en associant de manière linéaire et covalente un acide aminé à chaque nouveau codon de l'ARNm, une chaîne polypeptidique sera synthétisée en traduisant fidèlement l'information portée en amont par le gène présent sur l'ADN.

Dans les cellules vivantes, la biosynthèse des protéines a lieu au niveau des ribosomes. Ces macromolécules ribonucléoprotéiques sont des complexes enzymatiques qui catalysent les étapes-clés de la synthèse protéique. La traduction étant un processus complexe, un grand nombre de protéines, appelées facteurs de traduction, vont venir seconder le ribosome dans sa tâche. Comme toute réaction de polymérisation, la traduction peut être divisée en trois phases : le démarrage, l'élongation et la terminaison. Le recyclage du ribosome, en fin de traduction, peut éventuellement être considéré comme une quatrième étape.

La phase de démarrage de la traduction est déterminante pour la synthèse peptidique car elle fixe le codon de départ sur l'ARNm, et par conséquent, le cadre de lecture du polynucléotide. Il s'agit donc d'une étape majeure de l'ensemble du cycle cellulaire, de laquelle vont dépendre la fidélité, l'efficacité et la régulation de l'expression des gènes. De manière universelle, la biosynthèse des protéines de tous les êtres vivants débute par l'incorporation d'une méthionine. Toute la difficulté de cette étape de démarrage va donc résider dans la capacité du complexe de démarrage à reconnaître le codon AUG codant pour la méthionine initiatrice sur l'ARNm. Pour réaliser cette reconnaissance rigoureuse, le ribosome va devoir faire appel à un ARNt unique dans la cellule : l'ARNt initiateur méthionylé. Un jeu complexe de facteurs de démarrage va en outre intervenir pour permettre cette reconnaissance.

Si le mécanisme général de la traduction est globalement conservé dans les différents domaines du monde vivant, un regard plus approfondi sur les étapes successives met en évidence des différences significatives entre les traductions bactérienne, eucaryotique et archéenne, tant au niveau du nombre et de la nature des facteurs protéiques mis en jeu que du déroulement même du processus. Un exemple de cette divergence correspond au facteur protéique chargé de recruter spécifiquement l'ARNt initiateur méthionylé sur la petite sous-unité ribosomale. Chez les Eucaryotes et les Archées, ce facteur est un hétérotrimère respectivement nommé eIF2 et aIF2. Malgré son rôle crucial dans l'étape de démarrage, il ne possède pas d'orthologue bactérien. En effet, chez les Bactéries, l'ARNt initiateur méthionylé est modifié par une réaction de N-formylation et un facteur monomérique, nommé IF2, est responsable de sa prise en charge sur le ribosome. De manière surprenante, les séquences de génomes ont finalement révélé la présence, chez les Eucaryotes et chez les Archées, d'un orthologue d'IF2 bactérien nommé e/aIF5B (voir **Tableau 1.26**). Ainsi, chez ces organismes coexistent le facteur hétérotrimérique e/aIF2 et le monomère e/aIF5B. La compréhension du mécanisme d'action et du rôle précis d'e/aIF2 vis-à-vis de l'ARNt initiateur constitue donc un enjeu majeur à la fois d'un point de vue purement mécanistique mais également pour étudier les relations évolutives entre les 3 domaines du monde vivant.

Dans cette introduction, nous allons nous intéresser à la phase de démarrage de la traduction. Nous avons choisi de décrire le processus en mettant en lumière les ressemblances et les divergences existant au sein des trois domaines du monde vivant. Dans un second temps, nous nous recentrerons sur le démarrage de la traduction eucaryotique et archéenne, et nous étudierons en détail le rôle du facteur de démarrage e/aIF2, qui constitue le sujet principal de ce travail de thèse.

## **Partie A : Le démarrage de la traduction bactérienne.**

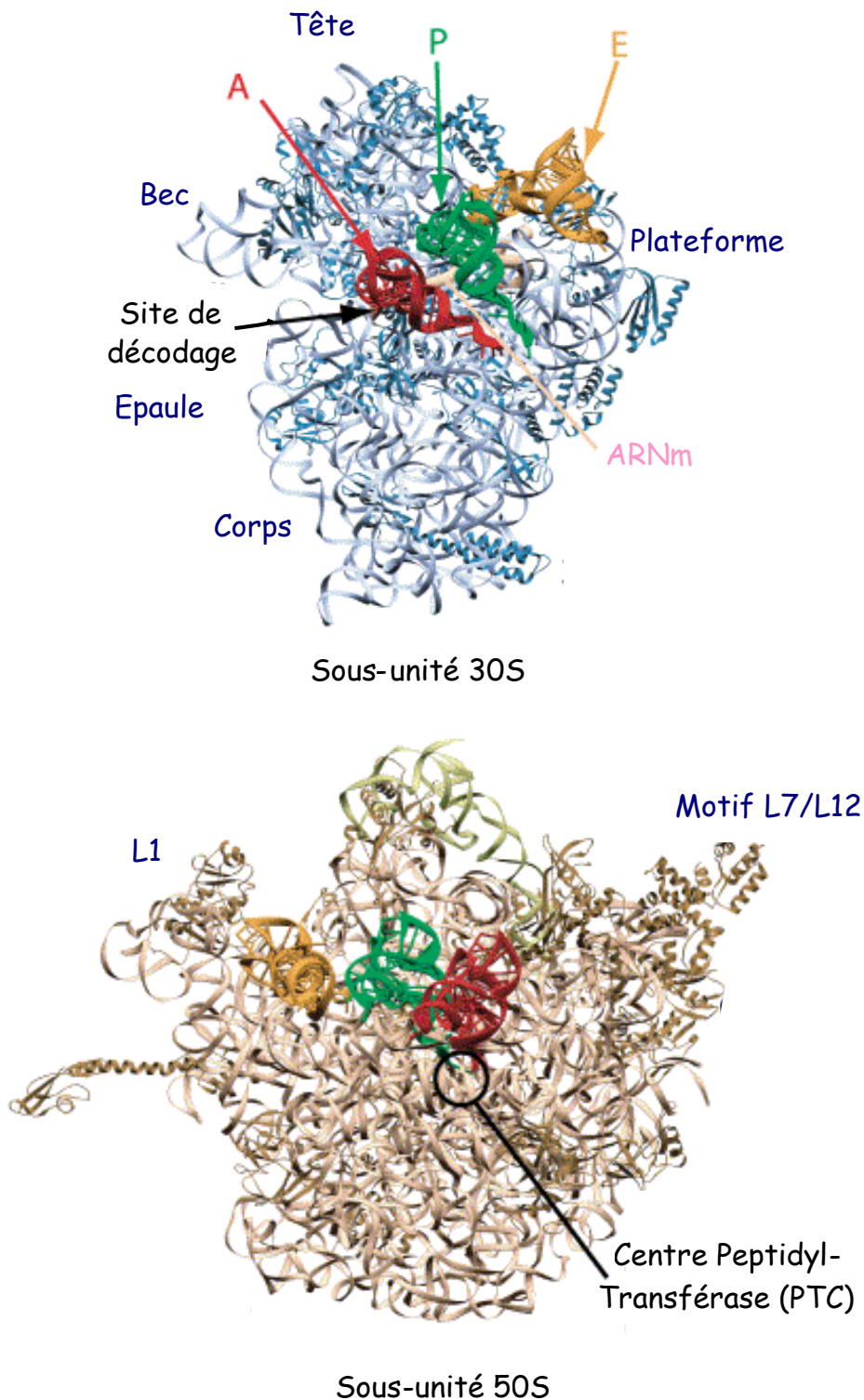
Le démarrage de la traduction est un processus complexe et hautement régulé qui implique des composants à la fois protéiques et ribonucléiques, à l'instar du ribosome qui accueille toute la phase de traduction. Chez les Bactéries, seulement trois facteurs protéiques interviennent dans l'étape de démarrage (IF1, IF2 et IF3), tandis qu'une douzaine de facteurs est nécessaire dans le système eucaryotique. Cette phase correspond à l'étape limitante du processus traductionnel : l'assemblage du ribosome sur l'ARNm nécessite quelques secondes et le ribosome traduit ensuite l'ARNm à une vitesse approximative de 12 acides aminés par seconde (Kennell and Riezman, 1977).

Au cours de ces dernières années, de nombreux progrès ont été faits dans la compréhension de ces mécanismes, en particulier grâce aux avancées spectaculaires réalisées dans le domaine de l'étude structurale des complexes macromoléculaires. Les efforts conjugués de la cristallographie par diffraction des rayons X et de la cryo-microscopie électronique ont permis d'obtenir des représentations précises de ces assemblages. Si la dynamique même du processus soulève encore de nombreuses interrogations, la mise en lumière par ces structures des interactions entre les différents partenaires a permis de mieux comprendre le déroulement de cette étape et une abondante documentation est désormais disponible sur le sujet (Sonenberg *et al.*, 2000 ; Gualerzi *et al.*, 2000 ; Boelens and Gualerzi, 2002 ; Laursen *et al.*, 2005 ; pour revue). Dans cette introduction, une description générale du mécanisme du démarrage de la traduction sera d'abord présentée. Puis les différentes étapes seront reprises en détail à travers l'étude des relations structure-fonction des composants de la machinerie biosynthétique.

### **1) Les étapes du mécanisme.**

Le ribosome bactérien (70S) est composé d'une grande sous-unité (50S) et d'une petite sous-unité (30S), elles-mêmes formées d'un assemblage de protéines et d'ARNs. La sous-unité 30S contient ainsi 21 protéines et un ARNr sédimentant à 16S. La sous-unité 50S est quant à elle composée de 34 protéines et de 2 ARNrs sédimentant respectivement à 5S et 23S. La structure tertiaire de l'ARN ribosomal est stabilisée par différents types d'interactions, notamment des interactions ARN-ARN, des interactions ARN-protéines et des pontages par des ions  $Mg^{2+}$  qui stabilisent les polynucléotides ribonucléiques phosphatés (Ban *et al.*, 2000 ; Nissen *et al.*, 2001). Le ribosome possède trois sites de liaison pour les molécules d'ARNt : le site A (aminoacyl) où se fixent les ARNts élongateurs, le site P

(peptidyl) qui fixe l'ARNt portant la chaîne polypeptidique naissante et accueille l'ARNt initiateur, et le site E (exit) où se retrouvent les ARNts déacylés avant leur éjection du ribosome. La **Figure 1.1** présente une vue générale des deux sous-unités ribosomales bactériennes.



**Figure 1.1 :** Vue générale de la petite et de la grande sous-unité du ribosome bactérien d'après Ramakrishnan, 2002. La face visible de chaque sous-unité correspond à la face qui interagit avec l'autre sous-unité au sein du ribosome 70S.

Le lancement d'un nouveau cycle de traduction nécessite la dissociation préalable du complexe 70S en sous-unités séparées. Cette dissociation est déclenchée par le facteur de recyclage RRF (Ribosome Recycling Factor) et la translocase EF-G (Elongation Factor G). Cette étape requiert l'hydrolyse d'une molécule de GTP portée par EF-G. Cette dissociation produit une sous-unité 50S libre et une sous-unité 30S complexée à l'ARNm et au dernier ARNt déacylé bloqué au site ribosomal P. L'éjection de ce dernier ARNt de la sous-unité 30S nécessiterait l'intervention du facteur de démarrage IF3 ( Karimi *et al.*, 1999; Agrawal *et al.*, 2004) (**Etape 1** de la **Figure 1.2**).

Une fois le complexe de post-terminaison dissocié, la réassociation 30S – 50S est empêchée grâce au rôle anti-associateur du facteur IF3 qui se fixe sur la petite sous-unité ribosomale. Le facteur IF1 vient stimuler l'activité anti-associatrice d'IF3 en se fixant également sur la 30S. L'action conjuguée de ces deux facteurs de démarrage pourrait également permettre de faciliter la dissociation de complexes 70S reformés prématurément.

Après dissociation du ribosome, IF2, l'ARNm et l'ARNt initiateur se fixent successivement à la sous-unité 30S dans un ordre qui reste à déterminer. L'ARNm est prépositionné sur le ribosome grâce à l'interaction d'une séquence dite de Shine-Dalgarno (SD), placée en amont de la séquence à traduire sur l'ARNm, avec une séquence nucléotidique qui lui est complémentaire sur l'ARNr 16S (ASD) (Shine and Dalgarno, 1974). Le positionnement précis de l'ARNm de manière à ce que le codon AUG initiateur se trouve au niveau du site ribosomal P requiert la présence des trois IFs. L'ordre exact dans lequel les différents partenaires du complexe de prédémarrage se fixent sur le ribosome n'étant pas certain, deux hypothèses sont à envisager. Soit les trois facteurs de démarrage sont déjà présents lors du recrutement de l'ARNm et par conséquent, le polynucléotide est directement placé dans une position optimale pour l'interaction codon-anticodon. Soit l'ARNm est prépositionné grâce à l'interaction SD-ASD au niveau du ribosome et le recrutement des IFs permet ensuite d'ajuster son positionnement de manière à ce que le codon initiateur se retrouve près du site P du ribosome (La Teana *et al.*, 1995).

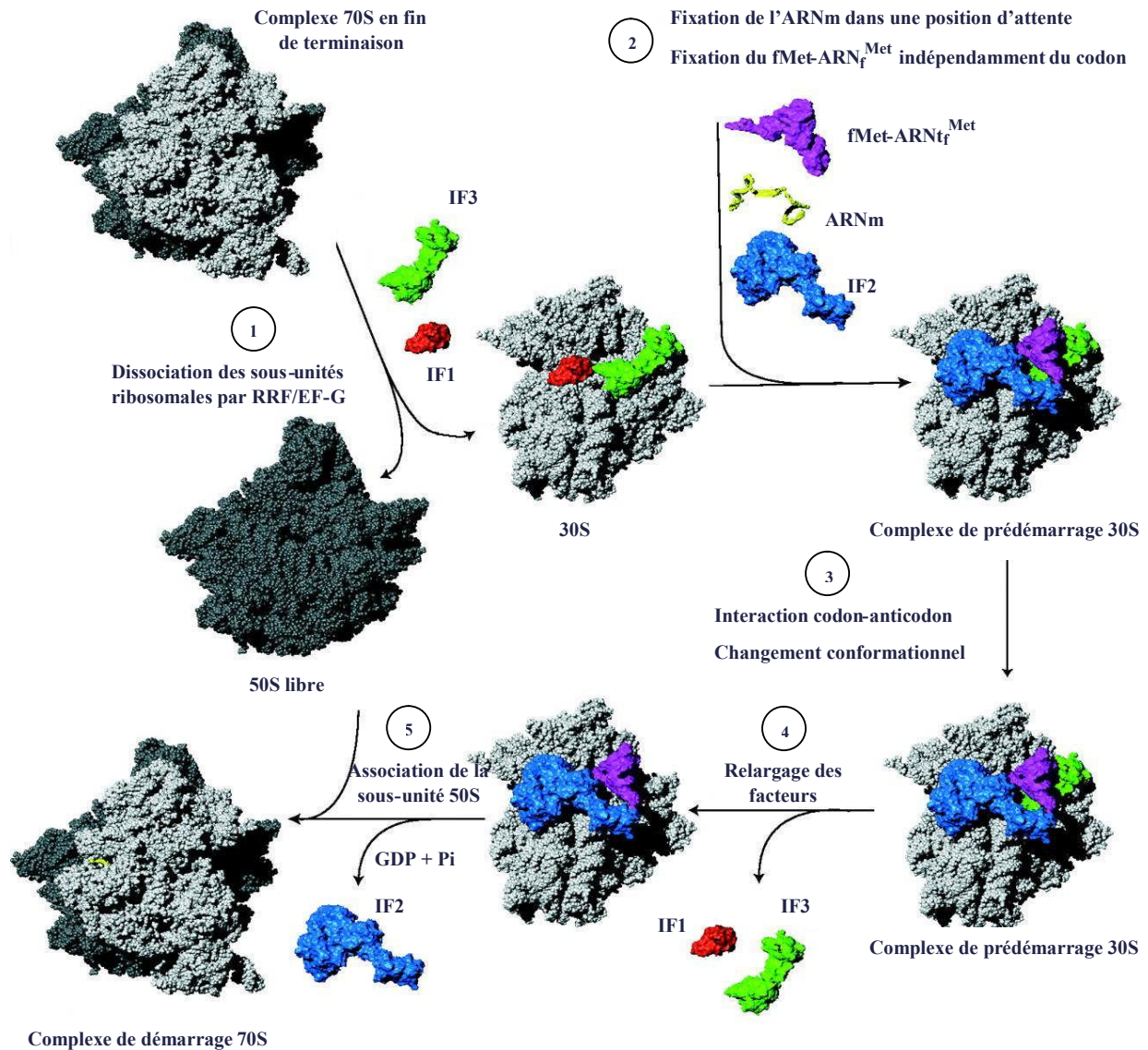
L'ARNt initiateur est également positionné au niveau du site P de la 30S selon un mode de liaison qui est tout d'abord indépendant de son interaction avec le codon de l'ARNm (Tomsic *et al.*, 2000). Le facteur IF1, qui se fixe au niveau du site A de la 30S, empêche ainsi le recrutement de l'ARNt initiateur au niveau du site A et il participe conséquemment à sa fixation au site P (Dahlquist and Puglisi, 2000 ; Carter *et al.*, 2001). IF3 stabilise également la liaison de l'ARNt initiateur au site P (Gualerzi *et al.*, 1977). Chez les Bactéries, l'ARNt initiateur porte une méthionine formylée (fMet-ARNt<sub>f</sub><sup>Met</sup>) et il est spécifiquement reconnu par le facteur IF2 avec lequel il interagit sur le ribosome (**Etape 2** de la **Figure 1.2**).

Le complexe de prédémarrage correspond donc à la 30S sur laquelle sont fixés les 3 facteurs de démarrage et l'ARNm dans une position d'attente avec le fMet-ARNt<sup>Met</sup> ancré sur la 30S de manière indépendante du codon initiateur. Ce complexe peu stable va alors subir un changement conformationnel qui permet l'interaction codon-anticodon entre l'ARNm et le fMet-ARNt<sup>Met</sup>, créant ainsi un complexe 30S stable (Gualerzi *et al.*, 1977 ; Pon and Gualerzi, 1984). La validité de cet appariement est ensuite vérifiée dans l'étape de décodage, en particulier grâce à IF3 qui est capable de dissocier un complexe pour lequel le codon de l'ARNm et l'anticodon de l'ARNt sont mésappariés (Hartz *et al.*, 1990) (**Étape 3** de la **Figure 1.2**).

Après relargage des facteurs IF1 et IF3, l'association de la sous-unité 50S pour former le complexe de démarrage 70S est stimulée par IF2. Le fMet-ARNt<sup>Met</sup> est alors ajusté au niveau du site P du ribosome, ce qui en fait désormais un substrat du centre peptidyltransférase (PTC) de la 50S, tandis que l'hydrolyse du GTP porté par IF2 en GDP et en Pi induit l'éjection d'IF2 du ribosome (**Étapes 4 et 5** de la **Figure 1.2**).

Le complexe de démarrage est alors prêt pour accueillir un aminoacyl-ARNt élongateur au niveau du site A qui arbore le second codon de l'ARNm. Cet ARNt est apporté par le facteur d'élongation EF1A sous forme d'un complexe ternaire avec une molécule de GTP. La première liaison peptidique entre cet acide aminé et la N-formyl-méthionine est alors formée au sein du PTC. Enfin, la première phase de translocation permet de faire glisser le peptidyl-ARNt ainsi formé vers le site ribosomal P, tandis que l'ARNt déacylé présent au site P est lui-même déplacé vers le site E. Ce glissement des molécules d'ARNt s'accompagne d'un mouvement parallèle de l'ARNm qui permet de maintenir la correspondance codon-anticodon pour chaque ARNt. Cette translocation assurée par EF-G libère le site A du ribosome qui est ainsi prêt pour entrer dans la phase d'élongation.

La **Figure 1.2** récapitule les différentes étapes de ce processus.

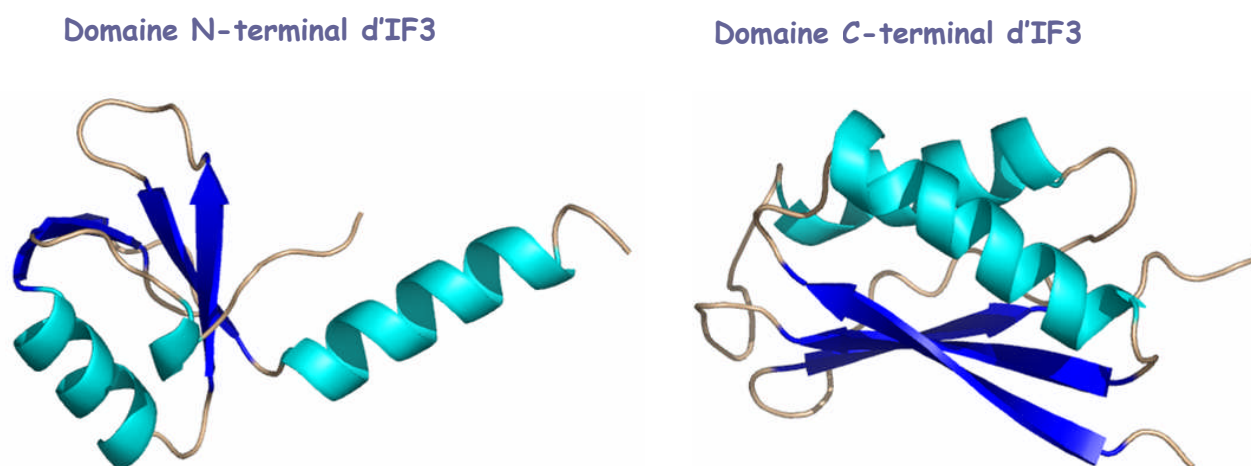


**Figure 1.2 :** Le démarrage de la traduction chez les Bactéries d'après Laursen *et al.*, 2005.

## 2) Formation du complexe de prédémarrage.

### 1. Activités anti-associatrices d'IF3 et d'IF1.

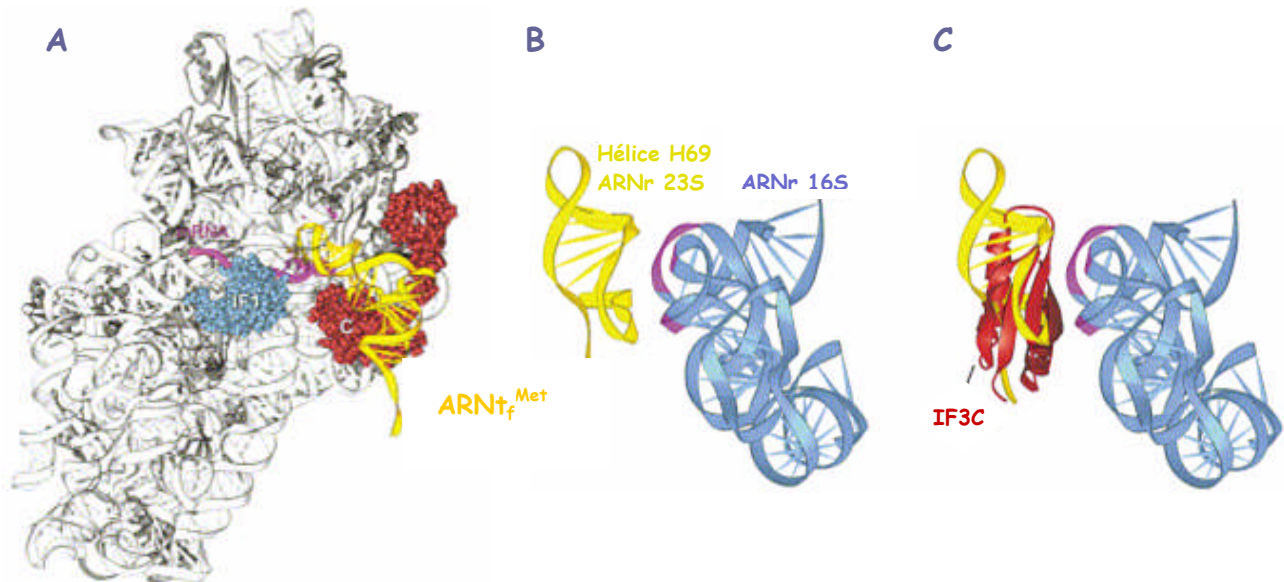
IF3 est une protéine de 20.4 kDa (soit 180 résidus) chez *E. coli*. Ce facteur comprend deux domaines de tailles voisines séparés par une région flexible d'environ 45 Å riche en lysines. La structure de ces deux domaines a été résolue par diffraction des rayons X et/ou spectroscopie RMN (Biou *et al.*, 1995 ; Garcia *et al.*, 1995). Les deux domaines sont représentés sur la **Figure 1.3**.



**Figure 1.3 :** Structure des domaines N- et C-terminaux d'IF3. A gauche, structure du domaine N-terminal d'IF3 de *B. stearothermophilus* (Biou *et al.*, 1995). A droite, structure du domaine C-terminal d'IF3 d'*E. coli* (Garcia *et al.*, 1995).

IF3 empêche la réassociation des sous-unités ribosomales en se liant à la sous-unité 30S (Grunberg-Manago *et al.*, 1975). Ses deux domaines présentent chacun des propriétés de liaison des ARNs et interagissent indépendamment l'un de l'autre avec la sous-unité 30S. Le domaine C-terminal (IF3C) est suffisant pour assurer les diverses fonctions biologiques du facteur. Cependant, en l'absence du domaine N-terminal (IF3N), son affinité pour la 30S est réduite (Petrelli *et al.*, 2001). IF3N permettrait donc de stabiliser l'association précaire entre IF3C et la plateforme de la sous-unité 30S en créant un second point de contact sur le ribosome. Des études cristallographiques et de marquage chimique ont permis de déterminer la position de ce facteur sur la sous-unité 30S. Les résultats obtenus par cryo-EM et par sonde avec des radicaux hydroxyles pointent une même zone d'interaction (McCutcheon *et al.*, 1999 ; Dallas and Noller, 2001) tandis que les données obtenues par diffraction des rayons X proposent un autre site de fixation d'IF3 (Pioletti *et al.*, 2001). Cependant ces résultats cristallographiques ont été remis en question car à cause de l'empilement du cristal, le site initial de fixation d'IF3 pourrait être masqué et IF3 se fixerait alors au niveau d'un site secondaire, du fait de ses propriétés générales de liaison des ARNs. La **Figure 1.4** illustre le modèle de positionnement déduit des expériences de marquage de l'ARNr 16S par sonde avec des radicaux hydroxyles. D'après ce modèle, les nucléotides protégés par IF3 et ceux protégés lors de la formation du complexe 70S correspondent à des zones qui se recoupent. Bien que la masse d'IF3 représente moins de 2% de celle de la 50S, la fixation du facteur sur la 30S empêcherait la 50S d'accéder à une large zone de la petite sous-unité ribosomale qui assure des contacts majeurs entre les deux sous-unités (Cate *et al.*, 1999 ; Yusupov *et al.*, 2001). Dans ce modèle, la position du domaine C-terminal d'IF3 coïncide avec celle de l'hélice 69 de l'ARNr 23S (**Figure 1.3**), cette hélice étant le principal constituant d'une zone pontant la

50S. Ainsi, IF3 pourrait empêcher la réassociation des sous-unités ribosomales en mimant cette hélice de l'ARNr 23S.

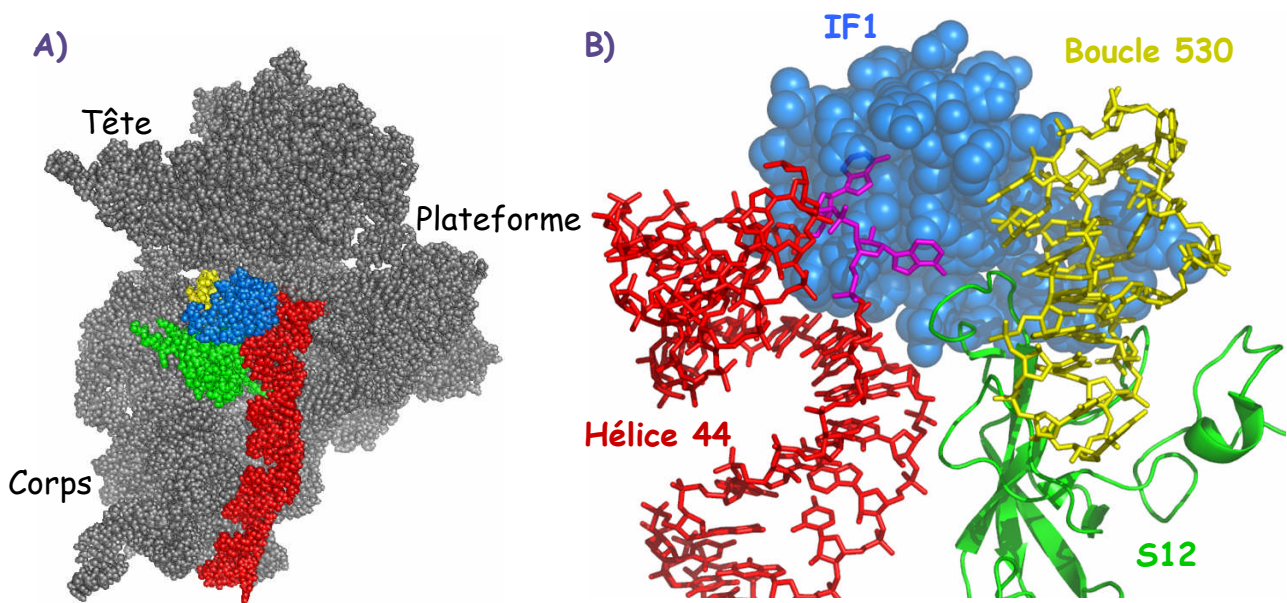


**Figure 1.4 :** Position du facteur IF3 sur la sous-unité ribosomale 30S relativement à IF1, à l'ARNm et à l'ARNt initiateur d'après Dallas and Noller, 2001. A) Modèle IF3-30S avec un ARNt initiateur au site P (en jaune) et IF1 (en bleu) localisé d'après la structure du complexe IF1-30S (Carter *et al.*, 2001). IF3 est représenté en rouge et l'ARNm en magenta. Le domaine C-terminal d'IF3 occupe la position de l'hélice 69 de l'ARNr 23S du complexe 70S. B) Interaction de l'hélice 69 (jaune) de l'ARNr 23S avec les hélices 23, 24 et 45 (bleues) de l'ARNr 16S. La zone d'interaction sur l'ARNr 16S est représentée en magenta. C) Superposition des positions d'IF3C (rouge) et de l'hélice 69 de l'ARNr 23S (jaune) sur la 30S.

Le caractère anti-associateur d'IF3 est renforcé par IF1. IF1 est le plus petit des 3 facteurs de démarrage de la traduction bactérienne. Il s'agit d'une protéine de 8.2 kDa chez *E. coli* et elle est composée d'un unique domaine qui présente un repliement classique de type OB-fold, caractéristique des protéines liant les acides nucléiques. Ce facteur est universellement conservé puisque l'on retrouve des homologues chez les Eucaryotes et chez les Archées (respectivement nommés eIF1A et aIF1A). Sa structure a été résolue par RMN (Sette *et al.*, 1997) (**Figure 1.18**).

IF1 augmente la vitesse d'association ou de dissociation des 30S avec les 50S mais il n'est pas lui-même responsable du déplacement dans un sens ou dans l'autre de l'équilibre réalisé entre les 70S et les sous-unités libres. Par contre, en présence d'IF3 qui bloque l'association des sous-unités ribosomales, IF1 va augmenter le taux de dissociation. Le positionnement d'IF1 sur la 30S a été étudié en détail, par diverses approches. La résolution de la structure de leur complexe (Carter *et al.*, 2001) a levé toutes les ambiguïtés. IF1 se lie au

niveau du site A de la petite sous-unité ribosomale. Lors de cette interaction, le facteur insère la boucle contenant les résidus 17 à 25 dans l'hélice 44 de l'ARNr 16S, ce qui repousse les nucléotides A1492 et A1493 du centre de décodage du site A vers l'extérieur (**Figure 1.5**). Le changement conformationnel ainsi amorcé se répercute au sein de la sous-unité sur une longue distance (environ 70 Å). L'état conformationnel obtenu pourrait correspondre à un état de transition de l'équilibre entre les formes associées et dissociées du ribosome, le facteur IF3 permettant alors de faire pencher cet équilibre en faveur de la forme non associée (Ramakrishnan, 2002).



**Figure 1.5 :** Position d'IF1 sur la sous-unité ribosomale 30S d'après Laursen *et al.*, 2005. **A)** Structure du complexe IF1-30S d'après Carter *et al.*, 2001. IF1 est représenté en bleu, l'hélice 44 de l'ARNr 16S en rouge, la boucle 530 en jaune et la protéine S12 en vert. L'hélice H44 (nucléotides A1492-A1493) et la boucle 530 (nucléotide G530) forment le centre de décodage du site ribosomal A. **B)** Vue rapprochée mettant en évidence les changements conformationnels induits par la fixation d'IF1. Les bases A1492 et A1493 sont indiquées en magenta.

Ainsi, *in vivo*, les complexes 70S post-terminaux sont dissociés par RRF et EF-G. Et la réassociation des sous-unités 30S et 50S ainsi libérées est bloquée par le facteur IF3 qui agit conjointement avec IF1. Toutefois, on ne peut exclure un rôle du couple IF3-IF1 dans la dissociation de complexes 70S qui se seraient réassociés après le processus de dissociation.

Dans tous les cas, à ce stade du démarrage de la traduction, la sous-unité 30S se retrouve liée aux facteurs IF1 et IF3. Elle doit encore s'associer avec IF2, l'ARNm et le fMet-ARNt<sup>Met</sup>. L'ordre dans lequel ces composés se fixent sur la 30S est cependant très controversé et notamment, il pourrait être totalement aléatoire.

## 2. Liaison de l'ARNm.

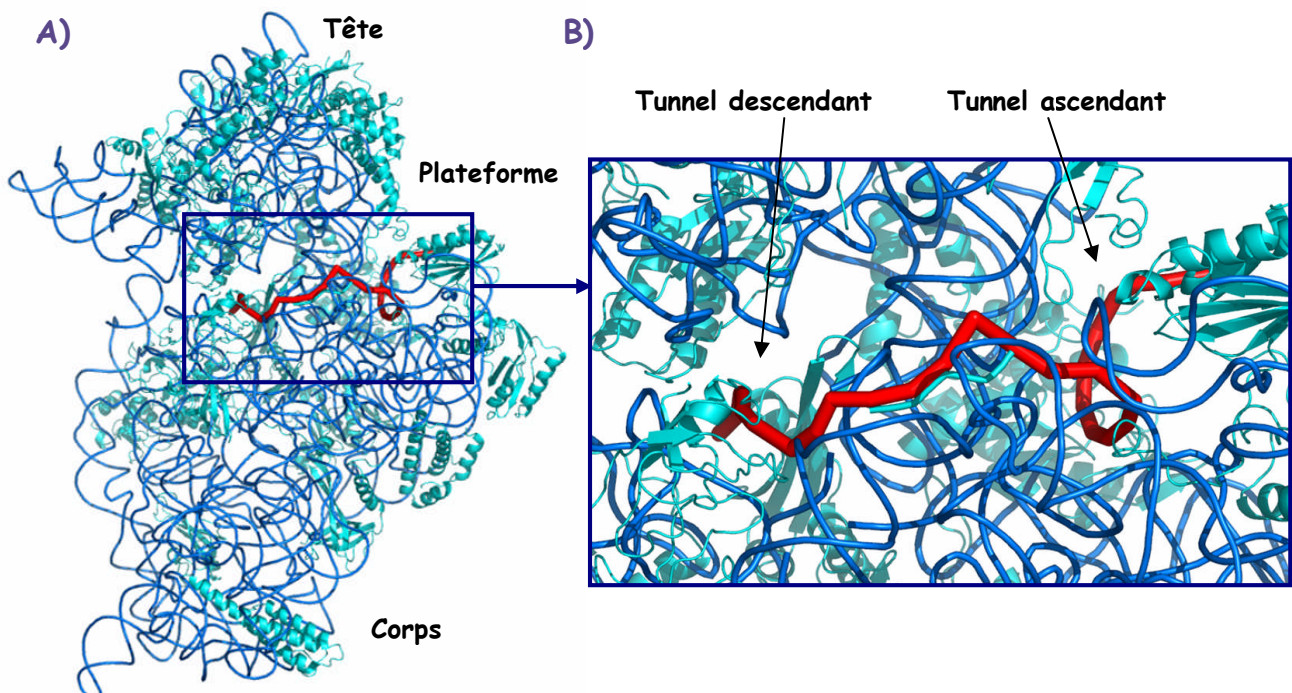
Pour former le bon complexe de démarrage, le codon initiateur doit être rigoureusement sélectionné sur l'ARNm et en particulier, il doit être distingué des autres codons AUG codant pour des méthionines internes de la protéine à traduire. De plus, chez les Bactéries, les ARNms sont polycistroniques, ce qui complique encore le travail de sélection. Chaque séquence à traduire possède en général ses propres signaux de démarrage. Le positionnement correct de l'ARNm sur le ribosome, de manière à sélectionner le bon codon de démarrage au niveau du site ribosomal P, est contrôlé par différentes interactions entre l'ARNm et le ribosome. La plus importante d'entre elles concerne une séquence de l'ARNm placée en amont de la séquence à traduire, sur la région 5' non traduite (5'UTR), et qui est appelée séquence de Shine-Dalgarno (SD). La séquence consensus des ARNms d'*E. coli* est une séquence GGAGG située  $7 \pm 2$  nucléotides en amont de l'AUG initiateur. Chez *E. coli*, le démarrage de la traduction peut également survenir sur un codon GUG ou UUG (respectivement 8% et 1% des cas) (Schneider *et al.*, 1986). Cependant, pour chacun de ces trois codons initiateurs possibles, c'est toujours une méthionine qui est incorporée par l'intermédiaire de l'ARNt initiateur N-formyl-méthionylé. Cette séquence SD va s'apparier avec une séquence complémentaire présente à l'extrémité 3' de l'ARNr 16S de la petite sous-unité 30S, appelée séquence anti-Shine Dalgarno (ASD) (Shine and Dalgarno, 1974).

Une autre interaction survient entre la protéine ribosomale S1 et une région de l'ARNm riche en pyrimidines qui se trouve en amont de la séquence SD. Cette région se comporte comme un site de reconnaissance du ribosome (Boni *et al.*, 1991) et une interaction directe a été mise en évidence par cryo-EM (Sengupta *et al.*, 2001).

Enfin, une région en aval du codon de démarrage appelée DB (Downstream Box) est rencontrée sur certains ARNms et elle présente une forte complémentarité avec les bases 1469 à 1483 de l'hélice 44 de l'ARNr 16S. Cette interaction DB-ADB serait donc de même nature que l'interaction SD-ASD (Sprengart *et al.*, 1990). Cependant aucune preuve génétique ou biochimique n'a pu mettre en évidence la nécessité de cette interaction pour le recrutement de l'ARNm par le ribosome. De plus, la structure de la sous-unité 30S de *T. thermophilus* suggère que l'épaule de la 30S s'insère entre les deux zones potentielles d'interaction (Wimberly *et al.*, 2000).

L'ARNm est d'abord prépositionné sur le ribosome au niveau d'un site d'attente (« standby site ») puis il est ajusté au niveau du site P par les différents facteurs de démarrage, en particulier IF3 (La Teana *et al.*, 1995).

Le site de liaison du ribosome sur l'ARNm (RBS pour Ribosomal Binding Site) s'étend sur une trentaine de nucléotides. Le positionnement de l'ARNm sur le ribosome a été établi par cryo-EM (Frank *et al.*, 1995) et plus récemment par diffraction des rayons X (Yusupova *et al.*, 2001). L'ARNm s'enroule autour du « cou » de la sous-unité 30S à mesure qu'il progresse au sein des tunnels ascendant et descendant (Culver, 2001). Un mouvement de loquet entre la tête et le corps de la 30S qui survient lors de l'activation du ribosome est à l'origine de ces deux tunnels. Enfin, pour progresser dans ces tunnels, l'ARNm doit être linéaire. Ainsi, sa liaison à la 30S *via* la séquence Shine-Dalgarno entraîne la disparition des structures secondaires de l'ARNm. Ce dernier pourrait être déroulé par des ARNm hélicases avant son entrée dans le tunnel descendant car une hélice d'ARN est trop large pour passer dans le tunnel (Olsthoorn *et al.*, 1995). La **Figure 1.6** représente le cheminement de l'ARNm sur la 30S.



**Figure 1.6 :** ARN messager lié à la sous-unité 30S d'après Yusupova *et al.*, 2001. **A)** Un ARNm de 36 nucléotides, représenté en rouge, est fixé sur la 30S (ARN ribosomal en bleu et protéines en cyan). **B)** Vue rapprochée. Les tunnels ascendant (upstream) et descendant (downstream) sont indiqués par des flèches.

### 3. Liaison de l'ARNt.

#### • Reconnaissance par IF2.

L'ARNt initiateur est également recruté par le ribosome. Il est spécifiquement reconnu par le facteur de démarrage IF2. Par analogie avec le démarrage eucaryotique et la phase d'élongation, un modèle initial proposait que le fMet-ARNt<sub>f</sub><sup>Met</sup> soit acheminé sur le ribosome

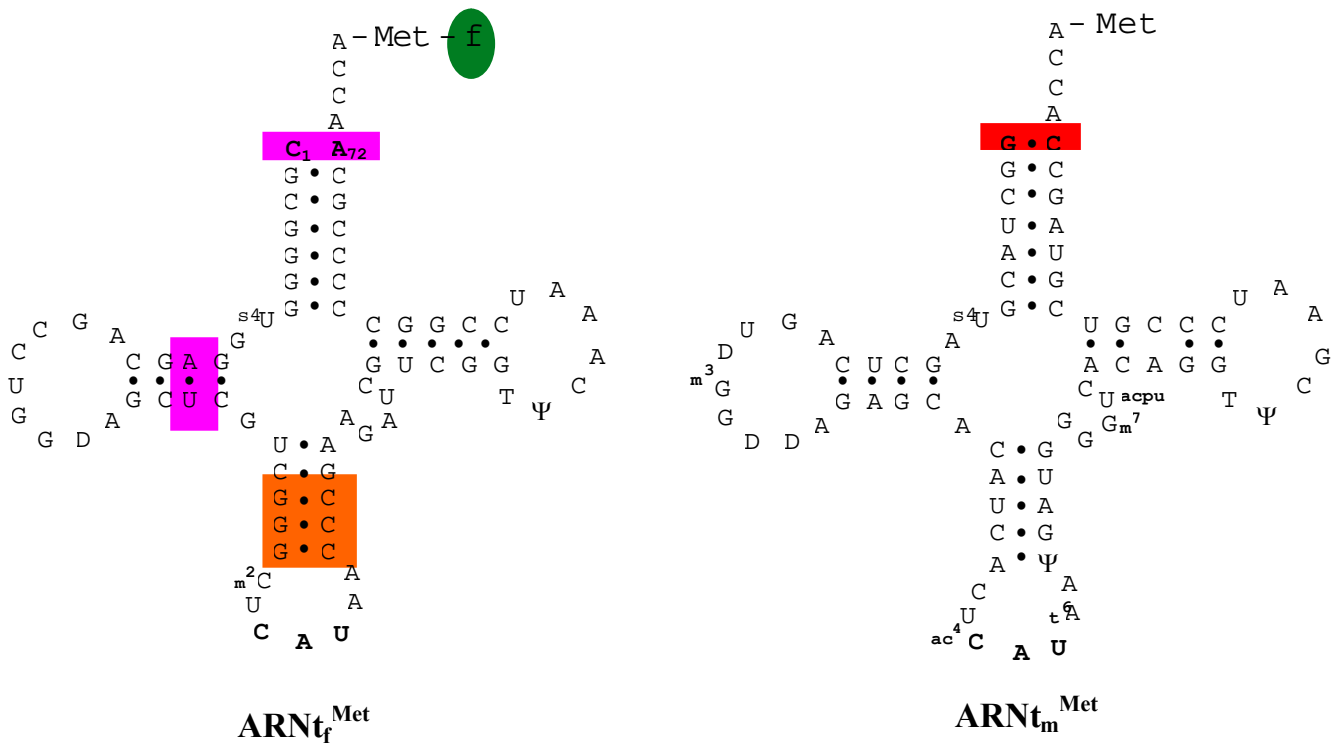
sous forme d'un complexe avec le facteur IF2 (tout comme les complexes ternaires eIF2:Met-ARNt<sub>f</sub><sup>Met</sup>:GTP et EF1A:aa-ARNt<sub>i</sub><sup>aa</sup>:GTP). Un complexe binaire entre IF2 et le fMet-ARNt<sub>f</sub><sup>Met</sup> est effectivement formé *in vitro*. Par contre, cette interaction est indépendante de l'état nucléotidique du facteur de démarrage (forme libre, GDP ou GTP). Ce complexe présente cependant un Kd faible (1,8 µM ; Petersen *et al.*, 1979). Un autre modèle a finalement été proposé selon lequel l'interaction entre IF2 et l'ARNt initiateur a lieu sur la sous-unité 30S (Canonaco *et al.*, 1986). Du fait de la très forte affinité d'IF2 pour la 30S (de l'ordre du nM), un complexe IF2-30S serait ainsi préformé et il sélectionnerait ensuite le fMet-ARNt<sub>f</sub><sup>Met</sup> parmi les autres ARNts cellulaires. Toutefois, certains résultats expérimentaux ne peuvent être expliqués par ce second modèle (Guillon *et al.*, 1996 ; Mayer *et al.*, 2003). Il est donc possible que les deux modes de recrutement coexistent au sein des cellules bactériennes.

L'ARNt initiateur qui sélectionne spécifiquement le codon initiateur AUG sur l'ARNm est un méthionyl-ARNt, aminoacylé par la méthionyl-ARNt synthétase (MetRS). Cependant, chez les Bactéries, une étape supplémentaire permet d'augmenter la discrimination positive de l'ARNt initiateur vis-à-vis des autres ARNts dans le processus de démarrage : la méthionine liée à l'ARNt initiateur est en effet N-formylée par la méthionyl-ARNt transformylase (MTF). Cette enzyme catalyse le transfert d'un groupe formyl du N<sup>10</sup>-formyltetrahydrofolate sur le groupe α-amine de la méthionine pour former le fMet-ARNt<sub>f</sub><sup>Met</sup>. Pour ce faire, elle doit donc distinguer le Met-ARNt initiateur du Met-ARNt élongateur. Ces deux ARNts présentent des caractéristiques nucléotidiques spécifiques de leur rôle dans la traduction (**Figure 1.7**).

La paire de bases 1-72, qui est non appariée dans le cas de l'ARNt initiateur, est un déterminant essentiel pour sa prise en charge spécifique par la formylase. Ce degré de liberté supplémentaire dans la chaîne polynucléotidique permet à la tige acceptrice de l'ARNt d'avoir une extrémité 3' mono brin suffisamment longue pour atteindre le site actif de la formylase, ce qui n'est pas le cas de l'ARNt élongateur du fait de l'appariement G-C en position 1-72 (Guillon *et al.*, 1992 ; Schmitt *et al.*, 1998).

De la même manière, l'absence d'appariement pour les positions 1 et 72 est également déterminante pour éviter l'interaction du fMet-ARNt<sub>f</sub><sup>Met</sup> avec la peptidyl-ARNt hydrolase (PTH). En effet, durant l'élongation, les ARNts présents au site P du ribosome et portant la chaîne polypeptidique naissante peuvent se dissocier malencontreusement du complexe 70S. Ces peptidyl-ARNts sont alors inutilisables par les cellules et ils doivent par conséquent être recyclés. C'est la PTH qui assure le clivage du polypeptide estérifié à l'extrémité de l'ARNt. L'action de la PTH s'exerce également sur les ARNts aminoacylés N-bloqués. Ainsi, le fMet-ARNt<sub>f</sub><sup>Met</sup> représente une cible potentielle pour cette enzyme. Des études biochimiques ont

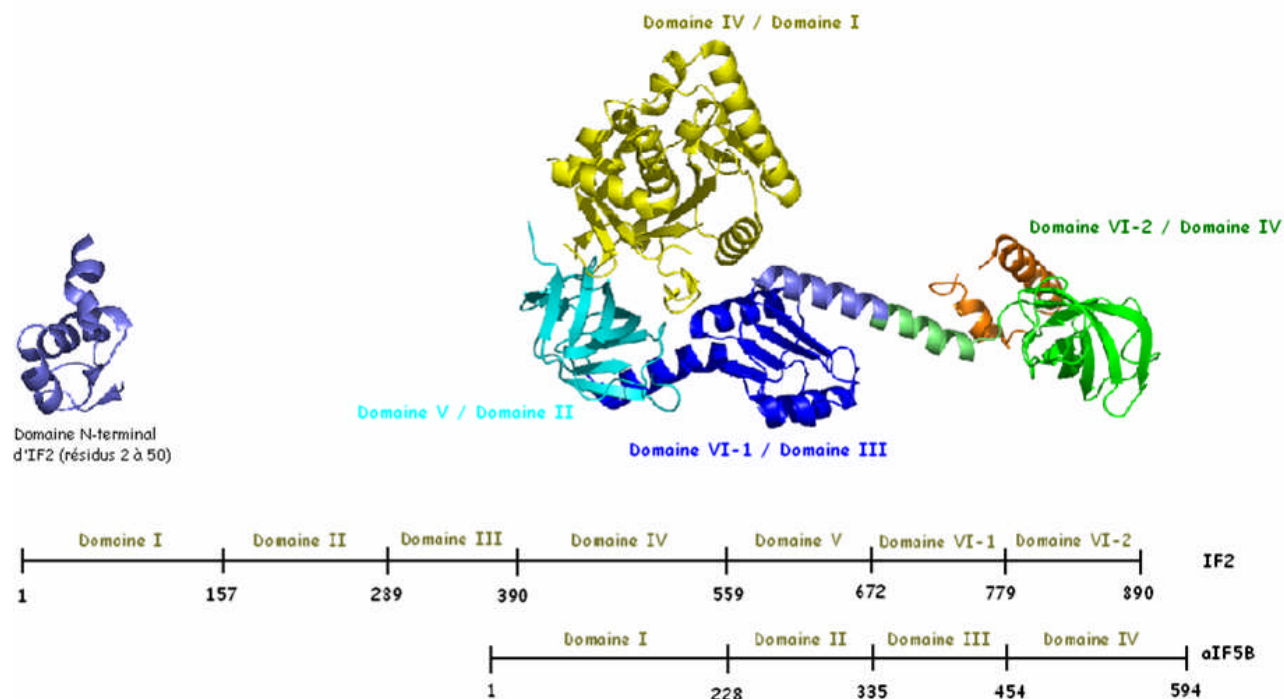
montré que la présence d'une paire de bases parfaitement appariée en position 1-72 est cruciale pour l'activité hydrolase de la PTH (Dutka *et al.*, 1993). L'enzyme interagit fortement avec le phosphate 5' du peptidyl-ARNt. Dans le cas de l'ARNt<sup>Met</sup>, le non-appariement des bases 1 et 72 rend le phosphate 5' très mobile et empêche donc sa reconnaissance par la PTH. La traduction peut ainsi avoir lieu.



**Figure 1.7 :** Caractéristiques nucléotidiques distinctives des méthionyl-ARNts initiateur et élongateur d'*E. coli*. Les régions importantes pour la reconnaissance spécifique de chaque ARNt par les différents partenaires de la traduction qui les prennent en charge sont indiquées comme suit : en magenta, régions reconnues par la MTF, les bases C<sub>1</sub> A<sub>72</sub> sont exclues par la PTH ; en rouge, appariement fort G<sub>1</sub>-C<sub>72</sub> exclu par la MTF ; en orange, déterminants de la tige anticodon du fMet-ARNt<sub>f</sub><sup>Met</sup> reconnu par IF3. Le groupe formyle responsable de la reconnaissance spécifique du fMet-ARNt<sub>f</sub><sup>Met</sup> par IF2 est encerclé en vert.

Le recrutement du fMet-ARNt<sub>f</sub><sup>Met</sup> est assuré par le facteur de démarrage IF2. C'est le plus volumineux des facteurs de démarrage bactériens et il est présent sous trois isoformes chez *E. coli* : IF2-1, IF2-2 et IF2-3 (97.3, 79.7 et 78.8 kDa respectivement). IF2 est composé de plusieurs domaines. Nous nous réfèrerons ici à la nomenclature concernant le facteur d'*E. coli* proposée par Mortensen *et al.*, 1998 et reprise sur la **Figure 1.8**. Ce facteur comprend une partie N-terminale (domaines I à III) de longueur variable et faiblement conservée. Au contraire, la partie C-terminale (domaines IV à VI) est universellement conservée puisque des homologues de cette région d'IF2 sont retrouvés chez les Eucaryotes et chez les Archées (respectivement nommés eIF5B et aIF5B) (Kyrpides and Woese, 1998 ; Sørensen *et al.*, 2001). La structure du domaine C-terminal d'IF2 a été résolue dans le cas de l'homologue

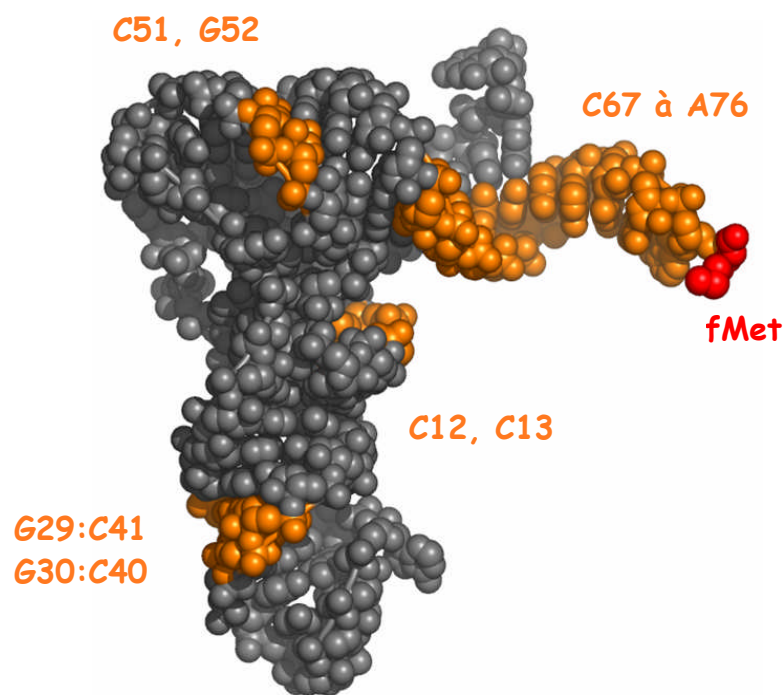
archéen  $\alpha$ IF5B de *M. thermoautotrophicum* (Roll-Mecak *et al.*, 2000) et plus récemment, la structure du domaine I d'IF2-1 bactérien a été résolue par RMN (Laursen *et al.*, 2003). La **Figure 1.8** représente l'organisation structurale d'IF2.



**Figure 1.8 :** Séquence primaire d'IF2 et organisation structurale du domaine I et des domaines IV à VI. A gauche, structure du domaine N-terminal d'IF2 d'*E. coli* (Laursen *et al.*, 2003). A droite, structure du facteur  $\alpha$ IF5B de *M. thermoautotrophicum* (Roll-Mecak *et al.*, 2000). L'homologue archéen possède 2 hélices additionnelles (en orange sur la structure d' $\alpha$ IF5B) à l'extrémité C-terminale du domaine VI par rapport au facteur IF2. La numérotation d'*E. coli* a été utilisée pour la séquence primaire d'IF2 et la numérotation de *M. thermoautotrophicum* a été utilisée pour la séquence primaire d' $\alpha$ IF5B.

Aucune activité catalytique majeure ne semble résider dans les trois domaines N-terminaux d'IF2. En effet, bien que le gène *InfB* codant pour IF2 chez *E. coli* soit indispensable à la croissance cellulaire, une forme tronquée pour laquelle les domaines I, II et III ont été enlevés est toujours capable d'assurer la croissance des bactéries à condition de surproduire la protéine mutante correspondante. Cette forme tronquée a d'ailleurs *in vitro* un taux d'activité voisin de celui d'IF2 natif dans les fonctions caractéristiques d'IF2 (Laalami *et al.*, 1991). Cependant, les domaines I et II d'IF2 pourraient faciliter la liaison du facteur sur le ribosome (Moreno *et al.*, 1998 ; Moreno *et al.*, 1999). Dans tous les cas, la variabilité de cette région N-terminale d'IF2 ne permet pas de généraliser les résultats obtenus. Au contraire, le domaine C-terminal semble assurer les principales fonctions du facteur. Il contient entre autre le domaine G, homologue à celui rencontré chez toutes les GTPases intervenant dans la traduction (EF1A, EFG, eIF2).

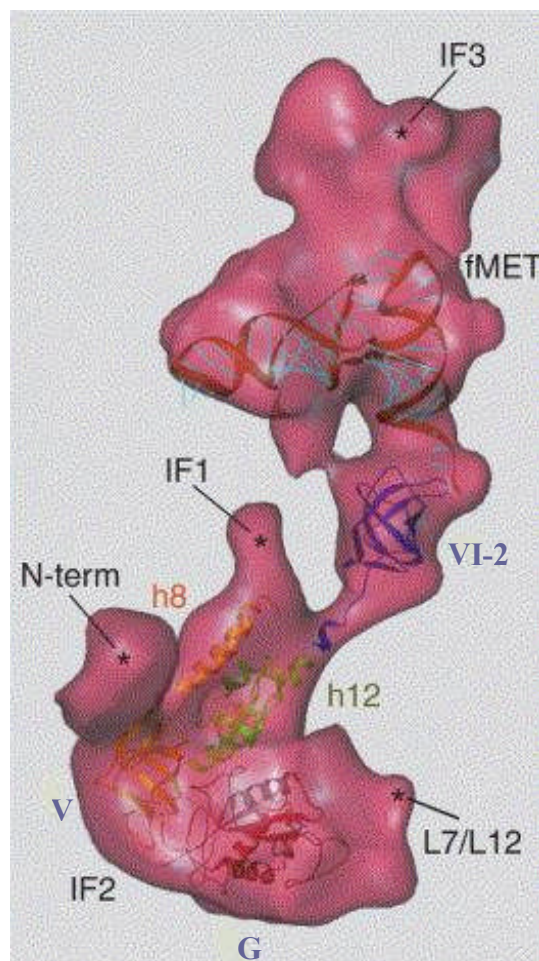
IF2 est responsable de la sélection du fMet-ARNt<sub>f</sub><sup>Met</sup>, probablement directement sur la 30S. La formation de ce complexe binaire se fait indépendamment du GTP (contrairement à la formation du complexe ternaire eIF2:Met-ARNt<sub>f</sub><sup>Met</sup>:GTP) (Sundari *et al.*, 1976 ; Pon *et al.*, 1985). Toutefois, lorsqu'IF2 se lie au fMet-ARNt<sub>f</sub><sup>Met</sup> sur la 30S, son affinité pour le GTP (mais pas celle pour le GDP) est considérablement augmentée et un complexe stœchiométrique peut être formé (Luchin *et al.*, 1999 ; Antoun *et al.*, 2003). Ce résultat suggère une fixation coopérative des trois partenaires sur la 30S (Wakao *et al.*, 1990). Les zones d'interaction entre IF2 et le fMet-ARNt<sub>f</sub><sup>Met</sup> ont été longuement étudiées. Ainsi, IF2 interagit principalement avec le groupement formyl du résidu méthionine de l'ARNt initiateur (Sundari *et al.*, 1976 ; Li *et al.*, 1996 ; Mayer *et al.*, 2003). La tige acceptrice pourrait également constituer un élément de reconnaissance pour IF2 (Mayer *et al.*, 2003). Deux autres sites de liaison ont également été identifiés par « footprinting » sur le bras T et sur la tige anticodon du polynucléotide (Wakao *et al.*, 1989 ; Yusupova *et al.*, 2001) (**Figure 1.9**).



**Figure 1.9 :** Zones d'interaction d'IF2 sur le fMet-ARNt<sub>f</sub><sup>Met</sup>. Les nucléotides reconnus par IF2 sont indiqués en orange et la formyl-méthionine est colorée en rouge.

Malgré ces divers contacts, c'est le groupe formyl qui constitue le déterminant majeur pour la reconnaissance spécifique du fMet-ARNt<sub>f</sub><sup>Met</sup>. En effet, la formylation empêche le fMet-ARNt<sub>f</sub><sup>Met</sup> d'être reconnu par EF1A et favorise au contraire sa reconnaissance par IF2 (Guillon *et al.*, 1993 ; Guillon *et al.*, 1996). La nature du résidu formylé semble être de seconde importance (Wu and RajBhandary, 1997).

Parallèlement, plusieurs sites d'interaction sur IF2 ont été proposés. Actuellement, l'hypothèse retenue est que le domaine C-terminal VI-2 d'IF2 contiendrait l'unique site de fixation de l'ARNt initiateur (Guenneugues *et al.*, 2000 ; Meunier *et al.*, 2000 ; Szkaradkiewicz *et al.*, 2000 ; Spurio *et al.*, 2000). Ce domaine possède une structure en tonneau  $\beta$  similaire à celle du domaine II de la GTPase EF1A (Nissen *et al.*, 1995). Chez EF1A, ce domaine II est impliqué dans la fixation de la tige acceptrice des ARNts élongateurs. Très récemment, une étude par cryo-EM a permis de visualiser un complexe de démarrage 70S bloqué avant l'étape d'hydrolyse du GTP (Allen *et al.*, 2005). Ces résultats confirment la reconnaissance du fMet-ARNt<sup>Met</sup> par le seul domaine C-terminal d'IF2 (**Figure 1.10**).



**Figure 1.10 :** Interprétation des cartes de densité du complexe de démarrage obtenues par cryo-EM et positionnement du complexe IF2-ARNt sur la 70S d'après Allen *et al.*, 2005.

Dans ce modèle, IF2 ne semble avoir qu'un seul point de contact au niveau de l'extrémité 3' de la tige acceptrice de l'ARNt. Les autres sites de liaison identifiés par footprinting sur l'ARNt ne sont pas expliqués. Il faut cependant noter que dans le contexte du complexe 70S, et donc après l'étape de décodage, des réarrangements conformationnels ont lieu. Ils permettent le relargage d'IF2 tandis que l'ARNt reste ancré sur le ribosome. Ce

modèle ne traduit donc probablement pas l'interaction exacte entre IF2 et le fMet-ARNt<sup>Met</sup> lors de leur fixation sur la 30S mais il souligne le rôle majeur joué par IF2-VI d'une part, et par la tige acceptrice de l'ARNt d'autre part, dans cette reconnaissance.

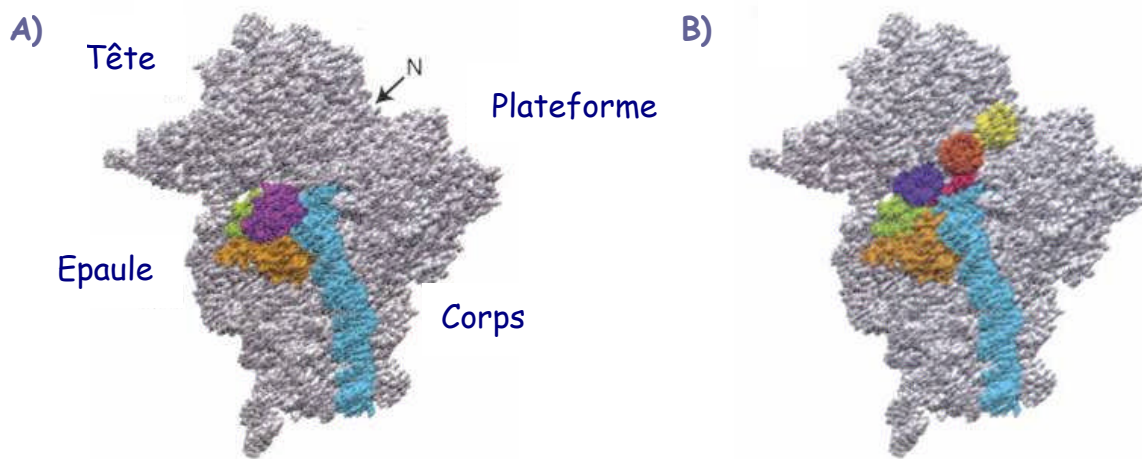
- **Site de fixation sur le ribosome.**

Dans le modèle actuel, l'ARNt initiateur reconnu par IF2 est acheminé au niveau du site P du ribosome, contrairement aux ARNts élongateurs qui arrivent d'abord au site A et subissent ensuite une translocation vers le site P après formation d'une nouvelle liaison peptidique. La tige anticodon de l'ARNt interagit avec l'ARNr 16S, tandis que la tige acceptrice et le bras T s'associeront avec l'ARNr 23S de la grande sous-unité ribosomale, la boucle du bras T interagissant alors avec la protéine ribosomale L23 (Yusupov *et al.*, 2001).

Deux paramètres sont à considérer pour expliquer pourquoi l'ARNt initiateur est dirigé vers le site P. Le premier concerne le facteur IF1. Comme l'ont révélé de nombreuses études de marquage et comme l'a confirmé la structure du complexe IF1-30S, IF1 se fixe au niveau du site A du ribosome, bloquant ainsi l'accès de ce site à l'ARNt (Carter *et al.*, 2001) (**Figure 1.11**). Le second concerne le facteur IF3. Comme le souligne la **Figure 1.4**, sur la sous-unité 30S, IF3 occupe une position qui recouvre le site E du ribosome. IF3 pourrait donc jouer un rôle similaire à celui d'IF1 au site A, en bloquant l'accès du site E à l'ARNt.

D'autre part, la présence d'IF1 sur la sous-unité 30S renforce la liaison du facteur IF2 sur le ribosome (Weiel and Hershey, 1982) et les deux facteurs ont ainsi une influence mutuelle sur leur liaison respective à la 30S. IF1 se lierait au niveau d'une région comprise entre les domaines III et VI d'IF2 (Moreno *et al.*, 1999). Cette hypothèse est cohérente avec la position d'IF2 sur le ribosome 70S (**Figure 1.10**) obtenue par cryo-EM. IF1 serait alors à proximité des domaines V et VI-2 d'IF2 ainsi que de la tige anticodon de l'ARNt.

L'obtention de nouvelles données de microscopie ou de cristallographie, mettant en œuvre un complexe IF2:fMet-ARNt<sup>Met</sup>:GTP:IF1:IF3:30S, permettra sans doute dans un avenir proche de déterminer définitivement la position de chaque partenaire sur la petite sous-unité ribosomale.



**Figure 1.11 :** Localisation d'IF1 sur la 30S d'après la structure du complexe (Carter *et al.*, 2001) et comparaison avec la position du site A. **A)** Position d'IF1 (en magenta) sur la 30S. L'hélice 44 de l'ARNr 16S (cyan), la boucle 530 (vert) et la protéine S12 (orange) sont également indiquées. **B)** Même vue. Le facteur IF1 a été retiré et la position des sites A (bleu marine), P (rouge) et E (jaune) a été indiquée. IF1 recouvre clairement le site A.

### **3) Interaction codon-anticodon et décodage.**

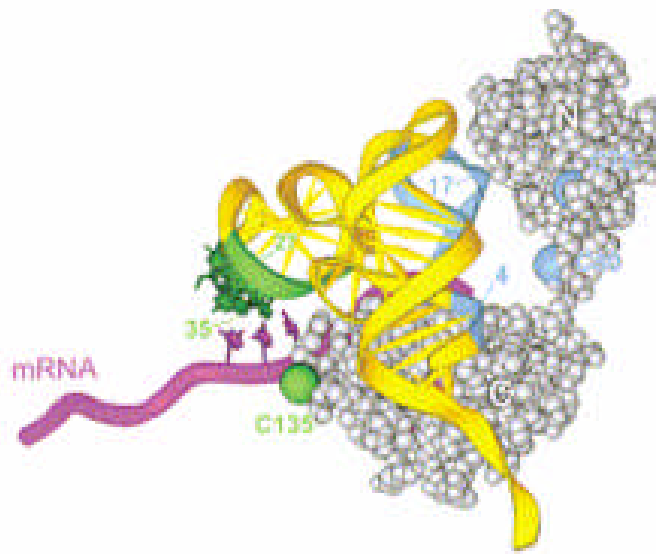
Lorsque tous les partenaires du démarrage de la traduction sont réunis sur la 30S, un réarrangement conformationnel va survenir pour créer un complexe 30S stable et permettre l'interaction codon-anticodon. Ce réarrangement entraîne l'ajustement de l'ARNm au site P du ribosome, sous l'influence des 3 facteurs de démarrage et particulièrement IF3 (La Teana *et al.*, 1995). L'ARNt initiateur se retrouve alors lié au ribosome selon un mode dépendant du codon de l'ARNm : l'interaction codon-anticodon peut avoir lieu. Les différentes structures de ribosomes bactériens ont montré une variation dans la position de l'extrémité 3' de l'ARNr 16S entre la structure de la 30S libre et celle de la 70S (Wimberly *et al.*, 2000 ; Yusupov *et al.*, 2001). Dallas et Noller (2001) proposent que l'interaction d'IF3 avec cette région de l'ARNr 16S repousse l'extrémité 3' de l'ARNr à l'opposé des sites ribosomiaux P et E, comme c'est le cas au sein de la 30S libre. Cette région deviendrait alors plus accessible et pourrait interagir avec la séquence de Shine-Dalgarno de l'ARNm, permettant un repositionnement du codon initiateur au plus près du site P.

La validité de l'appariement est ensuite vérifiée au cours de l'étape de décodage. Dans la phase d'élongation, le ribosome lui-même joue un rôle crucial dans le processus de décodage, grâce notamment à un changement conformationnel de l'ARNr 16S au niveau du site A lors de la fixation de l'ARNt élongateur, qui permet aux bases A1492, A1493 et G530 de venir en contact rapproché avec l'hélice codon-anticodon, vérifiant ainsi que l'appariement correct est réalisé. Pour la phase de démarrage, aucun changement conformationnel de

l'ARNr 16S ne survient lors de la fixation de l'ARNt initiateur au site P. Cependant, le nucléotide C1400 stabilise l'appariement codon-anticodon en venant au contact de la base 34 de l'ARNt (Yusupov *et al.*, 2001). Malgré cela, le ribosome ne semble pas exercer un contrôle très rigoureux sur la validité de l'interaction codon-anticodon et ce rôle reviendrait en fait majoritairement aux trois facteurs de démarrage (Hartz *et al.*, 1989).

IF2 joue un premier rôle déterminant, en permettant de sélectionner l'ARNt initiateur et il agit de concert avec le facteur IF1. Un rôle encore plus important est toutefois assigné au facteur IF3. Ce facteur semble en effet inspecter à la fois l'interaction codon-anticodon mais également la tige anticodon de l'ARNt<sup>Met</sup> (Hartz *et al.*, 1990). Ainsi, si l'une des trois paires de bases G-C caractéristiques de la tige anticodon de l'ARNt initiateur est modifiée, l'ARNt résultant n'est plus reconnu par IF3 et conséquemment, n'est plus sélectionné en tant qu'ARNt initiateur. Les complexes de démarrage qui présentent un aminoacyl-ARNt non initiateur ou un codon de démarrage différent des 3 codons autorisés (AUG, GUG et UUG, qui sont tous les trois reconnus par le fMet-ARNt<sup>Met</sup>) sont dissociés par IF3 (Hartz *et al.*, 1990). Concernant l'appariement codon-anticodon, deux hypothèses ont été envisagées pour le mode de reconnaissance par IF3. IF3 pourrait ainsi soit tester la complémentarité des 2 triplets de nucléotides, soit directement reconnaître la nature chimique des bases présentes au niveau du codon et de l'anticodon initiateur. Des expériences réalisées *in vivo* en présence d'ARNts initiateurs portant différents types d'anticodon sont en faveur de la première hypothèse (Meinzel *et al.*, 1999).

Le centre de décodage du site P se situe sur la petite sous-unité ribosomale, à l'interface entre la partie supérieure du « corps » et la partie inférieure de la « tête » de la 30S. Ce site est exclusivement composé d'ARNr et il contient notamment la partie supérieure de l'hélice 44 et les extrémités 3' et 5' de l'ARNr 16S (Schluzen *et al.*, 2000). Du fait de son rôle d'inspection de la tige et de la boucle anticodon de l'ARNt, IF3 pourrait interagir directement avec le polynucléotide. Le modèle de positionnement d'IF3 sur la 30S (**Figures 1.4 et 1.12**) tend d'ailleurs à valider l'hypothèse d'une interaction directe entre IF3C et la tige anticodon de l'ARNt. Une étude plus fine des contacts entre ces deux partenaires au niveau du site ribosomal P a alors été réalisée en utilisant un facteur IF3 modifié sur lequel des radicaux hydroxyles capables de cliver l'ARNt ont été incorporés. Ces expériences de clivage montrent qu'IF3C peut interagir avec l'anticodon ainsi que les nucléotides 26 à 29 de l'ARNt, ces derniers connectant le bras D à la tige anticodon. Par contre, les trois paires de bases G-C spécifiques de l'ARNt initiateur ne sont pas clivées par le facteur modifié. Or la nature de ces trois paires de bases est également validée par IF3 qui rejeterait un ARNt non-initiateur. Ainsi, un mécanisme indirect induit par IF3 pourrait également exister pour réaliser l'inspection de ces motifs nucléotidiques (Dallas and Noller, 2001).



**Figure 1.12 :** Vue rapprochée du modèle de positionnement d'IF3 sur la 30S d'après Dallas and Noller, 2001. L'ARNm est indiqué en magenta, l'ARNt initiateur lié au site P est en jaune. Les zones en vert et en bleu représentent les régions de l'ARNt coupées par les radicaux hydroxyles Fe(II) fixés sur les différentes cystéines d'IF3. L'ARNr et les protéines ribosomales ne sont pas représentées. La cystéine C135 d'IF3C clive les positions 35-37 de la boucle anticodon de l'ARNt. Par contre, les 3 paires de bases G-C de la tige anticodon ne sont pas directement modifiées.

Dans la structure du complexe 70S co-cristallisé avec un ARNt initiateur au site P, les nucléotides G1338 et A1339 de la 30S sont juxtaposés à la région de la tige anticodon (Yusupov *et al.*, 2001). Les expériences de cryo-EM (McCutcheon *et al.*, 1999) sur le complexe IF3-30S suggèrent qu'IF3 pourrait induire un mouvement de rabat de la « tête » de la 30S vers la plateforme. Ce mouvement aurait alors pour effet de placer les bases G1338 et A1339 en contact direct avec la tige anticodon de l'ARNt de manière à exercer un contrôle stérique de l'identité nucléotidique de la séquence. Ce modèle serait ainsi similaire à celui proposé pour le décodage des ARNts élongateurs au site A (Ogle *et al.*, 2003).

Enfin, une interaction directe entre IF2 et IF3 sur la 30S pourrait également avoir lieu et renforcer l'efficacité de l'étape de décodage.

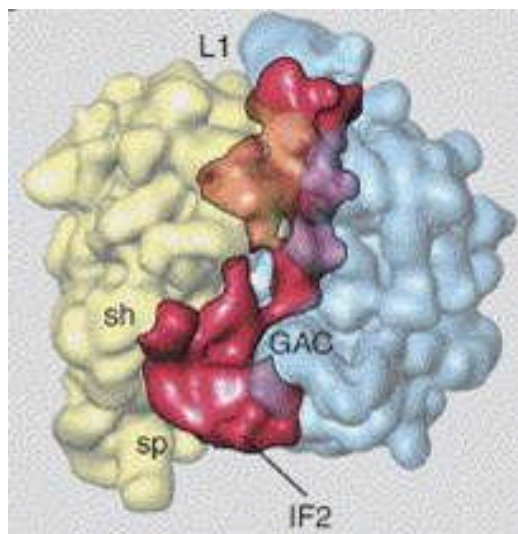
#### **4) Association des sous-unités ribosomales.**

Une fois l'appariement correct codon-anticodon réalisé, le complexe de prédémarrage peut entrer dans la dernière phase du processus : la fixation à la sous-unité 50S, pour former le complexe de démarrage final. Cette étape permet également l'ajustement final de l'ARNt au site P. Elle va nécessiter l'activation de la fonction GTPase d'IF2 qui n'a pas été sollicitée jusqu'à présent.

Le rôle exact de l'activité GTPase d'IF2 soulève encore aujourd'hui de nombreuses questions. Comme beaucoup de GTPases intervenant dans la traduction et/ou d'autres mécanismes biologiques cruciaux, IF2 ne possède pas d'activité GTPasique intrinsèque. Un activateur est donc requis pour hydrolyser le nucléotide porté par le facteur et pour IF2, c'est le ribosome qui joue ce rôle.

Contrairement au cas du trimère eIF2 eucaryotique, la présence de GTP, et par conséquent son hydrolyse, n'est pas requise pour qu'IF2 sélectionne précisément l'ARNt initiateur. Par contre, la forme IF2-GTP est nécessaire pour l'étape d'association des sous-unités ribosomales. De plus, la fixation de l'ARNt sur la 30S augmente significativement l'affinité d'IF2 pour le GTP. Il est donc probable que ce soit un complexe binaire IF2:GTP préformé qui recrute le fMet-ARNt<sup>Met</sup>, directement sur le ribosome ou bien antérieurement à l'ancrage d'IF2 sur la 30S. Dans tous les cas, le GTP est très certainement présent au sein du complexe de prédémarrage.

Le moment exact auquel survient l'hydrolyse du GTP n'est pas clairement déterminé. En effet, il a d'abord été supposé que cette hydrolyse était nécessaire pour favoriser la formation du complexe 70S. Plus récemment, diverses études ont suggéré que cette hydrolyse sert en réalité au relargage d'IF2 du ribosome (Luchin *et al.*, 1999 ; Antoun *et al.*, 2003). Ainsi, en présence de l'analogue non hydrolysable GDPNP, qui bloque l'activité GTPasique d'IF2, ou lors de l'utilisation de mutants d'IF2 qui n'hydrolysent plus le GTP, la formation des complexes 70S reste néanmoins effective. Par contre, la synthèse de la première liaison peptidique est bloquée. Ce résultat s'explique par la très forte affinité de la forme GTP-liée (ou GDPNP-liée) d'IF2 pour le ribosome : en l'absence d'hydrolyse du GTP, IF2 reste séquestré sur le complexe 70S et empêche ainsi la poursuite de la traduction. Ces résultats sont corroborés par la visualisation du facteur IF2 sur le complexe 70S (**Figure 1.13**) obtenue par cryo-EM (Frank *et al.*, 2005). Dans cette étude, le complexe de démarrage a été bloqué avant hydrolyse du GTP en utilisant à nouveau le GDPNP. L'absence d'hydrolyse du GTP ne semble donc pas empêcher l'association de la 30S avec la 50S.

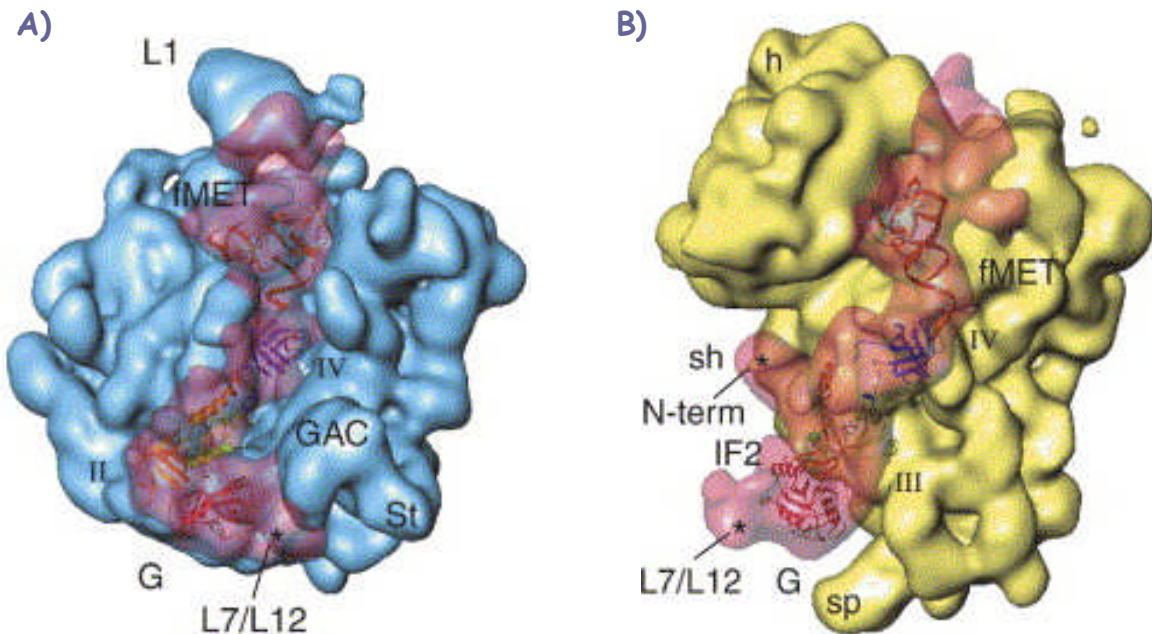


**Figure 1.13** : Structure du complexe 70S bloqué en présence d'IF2-GDPNP d'après Allen *et al.*, 2005. IF2 (rouge) interagit avec la 50S (bleu) au niveau du centre d'activation de la fonction GTPase (GAC pour GTPase Activating Center) et avec la 30S (jaune) sur une large zone à l'interface des 2 sous-unités, au niveau de l'épaule de la 30S.

Comme l'a révélé l'étude structurale de l'homologue aIF5B (Roll-Mecak *et al.*, 2000), la fixation du GTP sur IF2 entraîne un mouvement au sein du facteur. En effet, la liaison du nucléotide guanine provoque un léger déplacement de la région caractéristique Switch II du domaine G d'aIF5B. Ce déplacement est successivement transmis aux domaines II et III (qui correspondent respectivement aux domaines V et VI-1 d'IF2), puis il se répercute au niveau du domaine C-terminal (domaine VI-2 d'IF2) *via* le grand lien hélical. L'hélice et le domaine C-terminal sont ainsi entraînés dans un mouvement pendulaire qui induit une translation de 4.6 Å. L'effet se répercute donc sur une distance de plus de 90 Å au sein du facteur. La fixation du GTP sur IF2, après formation du complexe de prédémarrage, pourrait par conséquent placer le facteur dans une conformation active permettant l'assemblage ribosomal. Un repositionnement d'IF2 sur le ribosome a d'ailleurs été confirmé par l'étude structurale par cryo-EM (Frank *et al.*, 2005) par comparaison avec les études de clivage de l'ARNr 16S par IF2 au sein de la 30S (Marzi *et al.*, 2003). Cependant, la conformation adoptée par IF2:GDPNP au sein du complexe 70S (Frank *et al.*, 2005) diffère de celle observée pour le facteur libre (Roll-Mecak *et al.*, 2000) et suggère un réarrangement conformationnel beaucoup plus grand que celui observé dans le cristal lors de la transition GDP → GTP. Par conséquent, la fixation de la 50S sur le complexe de prédémarrage pourrait elle-même induire ce réarrangement au sein d'IF2.

L'étude par cryo-EM est particulièrement intéressante sur un second point : en permettant un positionnement d'IF2 sur le ribosome, elle met en lumière la présence du facteur à l'interface entre les deux sous-unités (**Figure 1.13**). De plus, le domaine G se situe

au niveau de la même région que les domaines G de deux autres GTPases de la traduction bactérienne : le facteur d'élongation EF1A et la translocase EF-G (**Figure 1.14**).



**Figure 1.14 :** Position d'IF2 sur chaque sous-unité du complexe 70S d'après Allen *et al.*, 2005. **A)** Sur la 30S. **B)** Sur la 50S.

Les domaines G de ces différentes GTPases présentent de nombreux points communs (voir discussion plus approfondie dans le **Chapitre II**), ce qui a conduit à imaginer un mécanisme commun d'activation de ces GTPases par le ribosome (Roll-Mecak *et al.*, 2000). Le fait que le domaine G d'IF2 se positionne au même endroit que le domaine G d'EF1A et d'EF-G tend à confirmer cette hypothèse même si des différences significatives subsistent très probablement pour chaque mécanisme. Les études menées sur le complexe ribosomal élongateur (Blanchard *et al.*, 2004) ont montré que la boucle P du domaine G d'EF1A, qui est responsable de la fixation du GTP, peut interagir avec la boucle sarcine-ricine (SRL) de l'ARNr 23S. Cette boucle SRL a ainsi été identifiée comme la zone activatrice de la fonction GTPase sur le ribosome. D'autres études structurales et de marquage croisé, sur le complexe de démarrage cette fois, ont suggéré que la partie C-terminale de la région L7/L12 s'intercalerait entre la protéine L11 et le domaine G d'IF2, résultats en accord avec les données cryo-EM. Cet élément L7/L12 est d'ailleurs indispensable à l'hydrolyse du GTP par EF1A et EF-G (Mohr *et al.*, 2002). Il pourrait ainsi induire un changement conformationnel sur le domaine G d'IF2 qui stimulerait son activité GTPasique.

L'échange  $\text{GTP} \rightarrow \text{GDP}$  induit alors probablement un réarrangement structural au sein d'IF2 en sens inverse de celui observé lors de la fixation du GTP. Ce mouvement aurait pour conséquence d'affaiblir l'interaction d'IF2 avec le ribosome et le facteur IF2-GDP, en conformation inactive, est ainsi éjecté du ribosome. Au cours de l'élongation, le facteur « inactif » EF1A-GDP doit être recyclé en un facteur actif EF1A-GTP car il s'associe préférentiellement avec le GDP. Ce recyclage fait appel à un facteur d'échange spécifique : EF1B. Chez les Bactéries, l'affinité d'IF2 avec le GDP n'est que 10 fois supérieure à son affinité pour le GTP. Et la concentration cellulaire en GTP est 10 fois plus élevée que celle du GDP. Un facteur d'échange n'est donc pas nécessaire pour obtenir un IF2 actif dans un nouveau cycle de démarrage (Pon *et al.*, 1985).

Le dernier point relatif à cette étape concerne le relargage des 2 autres facteurs de démarrage : IF1 et IF3. Celui-ci survient avant ou au plus tard, concomitamment à celui d'IF2. L'activité anti-associatrice d'IF3 laisse penser que le facteur quitte la 30S avant l'arrivée de la 50S ou bien que l'association 30S – 50S déclenche l'éjection du facteur. Dans tous les cas, le facteur IF3 devrait être absent du complexe 70S. Cependant, certaines données biochimiques récentes (Grigoriadou *et al.*, 2004) suggèrent qu'IF3 n'est relargué du ribosome qu'après l'hydrolyse du GTP. Enfin, le complexe de démarrage 70S bloqué avant hydrolyse du GTP visualisé par cryo-EM pourrait également contenir le facteur IF3, ainsi qu'IF1 (Allen *et al.*, 2005). Ainsi, l'hydrolyse du GTP pourrait être le déclencheur commun de l'éjection des trois facteurs de démarrage. Quelque soit l'ordre exact du déroulement de ces étapes, les 3 facteurs sont finalement relargués et il reste un complexe 70S libre de tout facteur, avec un ARNt initiateur ajusté au site P et prêt à entrer dans le cycle d'élongation.

## **Partie B : Le démarrage de la traduction eucaryotique.**

L'objectif général à atteindre lors de la phase de démarrage de la traduction - à savoir la formation d'un complexe de démarrage portant l'ARNt et l'AUG initiateurs au site P - est le même dans les trois domaines du monde vivant. De ce fait, certaines similitudes sont rencontrées dans les grandes lignes du déroulement des processus bactérien et eucaryotique. D'ailleurs, les facteurs de démarrage bactériens IF1 et IF2 possèdent des orthologues eucaryotiques et archéens : les facteurs e/aIF1A et e/aIF5B, respectivement (**Tableau 1.26**). De plus, comme nous le verrons dans cette partie, le facteur e/aIF1 présente une structure tridimensionnelle très proche de celle du domaine C-terminal d'IF3, c'est-à-dire le domaine fonctionnel du facteur bactérien. Des fonctions similaires à celles des facteurs bactériens sont ainsi assurées par ces homologues dans le démarrage de la traduction eucaryotique. Cependant, certaines divergences significatives existent entre les deux démarrages. Ainsi, l'ARNt initiateur méthionylé eucaryotique n'est pas formylé. D'autre part, chez les Eucaryotes, les ARNms ne possèdent pas de séquence SD permettant leur prépositionnement sur le ribosome. Enfin, une machinerie différente, passant par une large panoplie de facteurs (eIF1, eIF1A, eIF2, eIF2B, eIF3, eIF4A, eIF4B/eIF4H, eIF4E, eIF4G, eIF5, eIF5B, eIF6), est rencontrée dans le système eucaryotique.

Nous essaierons donc dans cette partie de décrire le processus de démarrage de la traduction eucaryotique en l'état actuel des connaissances, en mettant à nouveau en parallèle la relation entre la structure des composants de cette machinerie et leur fonction biologique au sein de l'étape (Kapp and Lorsch, 2004 ; Preiss and Hentze, 2003 ; Pestova and Hellen, 2000 ; Sonenberg and Dever, 1998 ; Pain, 1996 ; pour revues).

### **1) Les étapes du mécanisme.**

Tout comme le ribosome bactérien, le ribosome eucaryotique (80S) est composé d'une petite sous-unité (40S) et d'une grande sous-unité (60S). La sous-unité 40S contient un ARNr sédimentant à 18S et 33 protéines. La sous-unité 60S est formée de 3 ARNrs (5S, 5.8S et 28S) et de 49 protéines. Le ribosome eucaryotique contient ainsi un plus grand nombre de protéines que le ribosome bactérien et des insertions sont présentes dans les ARNrs 18S et 28S par rapport aux ARNrs bactériens 16S et 23S. Cependant, l'architecture de cet assemblage macromoléculaire présente des similitudes frappantes dans les trois domaines du monde vivant, en particulier au niveau de la cavité entre les deux sous-unités ribosomales qui va

accueillir les molécules d'ARNt et également au niveau de la localisation des centres peptidyltransférase (PTC) et de décodage (Spahn *et al.*, 2001).

Dans le système eucaryotique, la traduction débute également par la dissociation des sous-unités ribosomales. Néanmoins, les intervenants du recyclage du ribosome restent très mystérieux. En effet, il n'existe aucun orthologue du facteur bactérien RRF chez les Eucaryotes. Les facteurs eIF1A, eIF3 et eIF6 (et éventuellement eIF1, par comparaison avec IF3C) possèdent tous une activité anti-associatrice *in vitro*. Cependant, aucun résultat expérimental n'indique clairement si ces activités sont impliquées dans le recyclage du ribosome ou bien si elles servent simplement à empêcher une association prématurée des 40S et des 60S au cours du démarrage de la traduction. Ainsi, le mécanisme par lequel survient le recyclage du ribosome n'a toujours pas été clarifié (Kapp and Lorsch, 2004).

Le Met-ARNt<sup>Met</sup> est reconnu par le facteur eIF2, indépendamment du ribosome, et il est acheminé sous forme d'un complexe ternaire avec eIF2-GTP jusqu'à la 40S. Un complexe de prédémarrage 43S est alors formé. Celui-ci comprend, outre la 40S et le complexe ternaire, les facteurs eIF1, eIF1A, eIF5 et le facteur multimérique eIF3. Ce dernier facteur est le plus volumineux des facteurs de démarrage de la traduction eucaryotique. Il est composé de 5 sous-unités chez la levure et d'au moins 12 sous-unités chez l'homme. En formant un complexe avec les différents facteurs de démarrage (MFC pour Multi-Factor Complex), il permet de ponter les différents partenaires et de renforcer leur arrimage sur le ribosome (**Etape 1** de la **Figure 1.15**).

Le complexe de prédémarrage va alors se fixer sur l'ARNm. Les ARNms eucaryotiques ne possèdent pas de séquence de Shine-Dalgarno en amont de la séquence à traduire et ils ne peuvent donc pas être prépositionnés sur le ribosome comme le sont les ARNms bactériens. Un autre mode de liaison est par conséquent mis en place dans le système eucaryotique. En fait, le complexe 43S va se fixer à l'extrémité 5' de l'ARNm qui porte une coiffe m<sup>7</sup>G(5')ppp(5')N. La reconnaissance de cette extrémité 5' coiffée passe par un complexe protéique unique, propre au système eucaryotique : eIF4F (eIF4A, eIF4E, eIF4G). Ce complexe reconnaît également l'extrémité 3' polyadénylée de l'ARNm, *via* une interaction entre eIF4G et les protéines PAB (Poly-A Binding) qui se fixent sur la queue poly-A des ARNms eucaryotiques. Cette double reconnaissance entraîne une circularisation de l'ARNm à traduire. Le rôle de cette circularisation sera discuté en détails plus loin dans cette partie. Sous l'action du facteur eIF4B (et eIF4H chez les mammifères), qui stimule l'activité RNA-hélicase d'eIF4A, les structures secondaires de l'extrémité 5' de l'ARNm sont ensuite déroulées et le complexe 43S peut venir se fixer sur l'extrémité 5' de l'ARNm, grâce

notamment à l'interaction entre eIF3 et les facteurs eIF4G et eIF4B (**Etape 2** de la **Figure 1.15**).

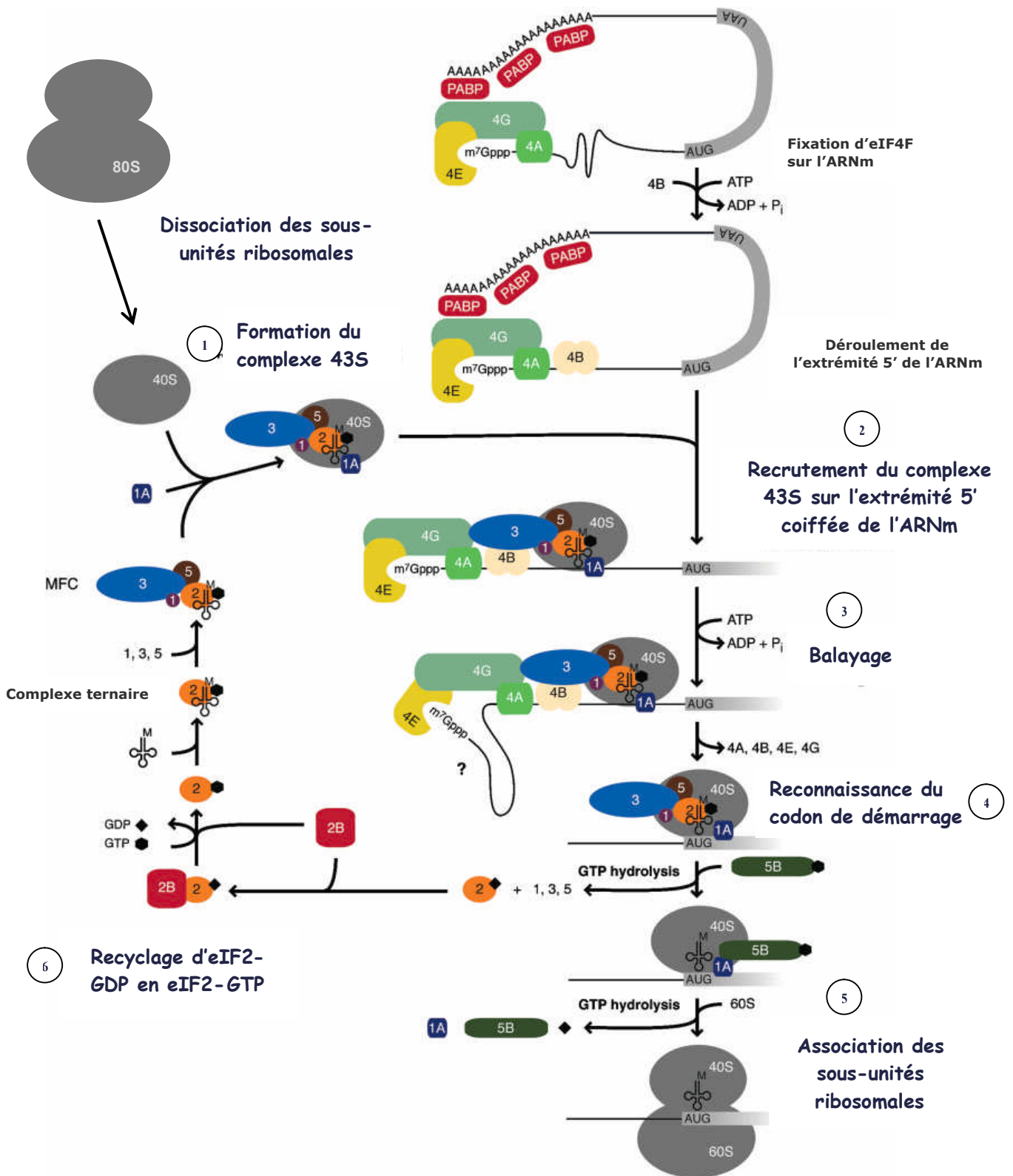
Le bon codon de démarrage est ensuite repéré grâce à un balayage de l'ARNm. En effet, le complexe ribosomal formé sur l'extrémité 5' de l'ARNm glisse le long de la séquence du polynucléotide jusqu'à rencontrer l'AUG initiateur. Les ARNms eucaryotiques sont généralement monocistroniques et l'AUG initiateur est le premier rencontré sur la séquence. Cependant, comme nous le verrons un peu plus loin, sa reconnaissance est loin d'être aussi simple. L'étape de balayage nécessite la présence des deux facteurs eIF1 et eIF1A (**Etape 3** de la **Figure 1.15**).

Lorsque l'AUG initiateur est reconnu sur l'ARNm, l'appariement codon-anticodon est vérifié au sein de la 40S. Plusieurs facteurs ont été impliqués dans cette sélection, en particulier les 3 sous-unités d'eIF2, eIF1 et eIF5. eIF3 pourrait également jouer un rôle dans cette reconnaissance. Lorsque l'appariement correct codon-anticodon est réalisé, l'hydrolyse de la molécule de GTP portée par eIF2 est déclenchée. Comme IF2, eIF2 ne possède pas d'activité GTPasique intrinsèque. Chez les Eucaryotes, un activateur existe pour faciliter l'hydrolyse du GTP sur l'hétérotrimère. Ce facteur qui joue le rôle de GAP (GTPase Activating Protein) est eIF5. D'autre part, la 40S est également requise pour cette fonction. Elle jouerait donc le rôle d'un second activateur. Dans le modèle actuel, l'hydrolyse du GTP par eIF2 entraîne son relargage du ribosome ainsi que celui des autres facteurs de démarrage à l'exception d'eIF1A. Simultanément, l'ARNt initiateur a été ajusté au niveau du site P (**Etape 4** de la **Figure 1.15**).

La dernière étape correspond à l'association des sous-unités ribosomales. Pour cela, un dernier facteur intervient, le facteur eIF5B, qui est l'homologue du facteur bactérien IF2. Le moment exact auquel survient le recrutement de ce facteur sur le ribosome n'est pas clairement défini. Cependant, comme son homologue bactérien, il favorise la formation GTP-dépendante du complexe de démarrage 80S. Après recrutement de la sous-unité 60S, la molécule de GTP portée par eIF5B est hydrolysée, ce qui entraîne l'éjection des derniers facteurs du ribosome : eIF5B et eIF1A. Et le complexe 80S est alors prêt pour entrer dans le cycle d'élongation (**Etape 5** de la **Figure 1.15**).

Le facteur eIF2-GDP est alors recyclé en eIF2-GTP par le facteur d'échange eIF2B (**Etape 6** de la **Figure 1.15**).

La **Figure 1.15** récapitule les différentes étapes décrites ci-dessus.



**Figure 1.15 :** Le démarrage de la traduction chez les Eucaryotes d'après Preiss and Hentze, 2003.

## 2) Formation du complexe de prédémarrage 43S.

### 1. Les partenaires protéiques du complexe de prédémarrage.

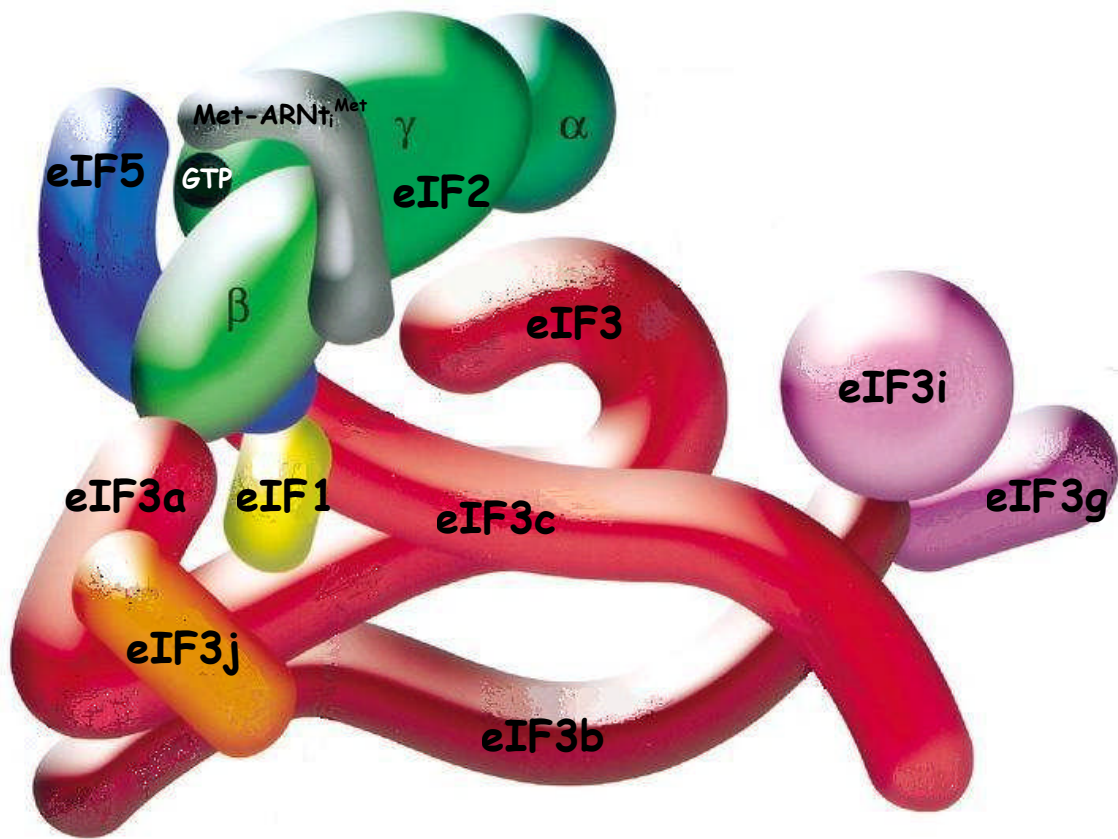
- **eIF3.**

eIF3 est le plus volumineux des facteurs de démarrage de la traduction eucaryotique. Le facteur mammifère contient 12 sous-unités, nommées eIF3a à eIF3l. Une treizième sous-unité, GA17, a été récemment identifiée (Unbehaun *et al.*, 2004). Chez la levure, seules 5 sous-unités forment le corps fonctionnel du facteur (TIF32/eIF3a, PRT1/eIF3b, NIP1/eIF3c, TIF34/eIF3i, TIF35/eIF3g). Une sixième sous-unité est également présente en quantités sous-stœchiométriques (HCR1/eIF3j). Par la suite, nous adopterons la nomenclature eIF3a-eIF3l pour nous référer aux sous-unités de tous les eIF3 eucaryotiques.

Le multimère eIF3 joue le rôle d'un immense adaptateur qui relie entre eux les différents facteurs de démarrage et autres composants de la machinerie traductionnelle. En effet, de multiples interactions ont été mises en évidence entre eIF3 et les facteurs eIF1, eIF1A, eIF2, eIF5, eIF4B et eIF4G (Vornlocher *et al.*, 1999 ; Phan *et al.*, 2001 ; Valasek *et al.*, 2002 ; Olsen *et al.*, 2003), ainsi que la 40S. De plus, eIF3 présente des domaines d'interaction avec de nombreuses protéines cellulaires telles que les protéines du cytosquelette ou du protéasome. Il pourrait donc servir de lien entre la machinerie traductionnelle et le reste de la cellule. A l'heure actuelle, seule la structure de la sous-unité k du facteur a été résolue chez l'homme (Wei *et al.*, 2004). eIF3k n'est cependant pas présente chez les levures. Il est donc probable que cette sous-unité n'ait qu'un rôle de régulation dans le processus de démarrage des Eucaryotes supérieurs. Par contre, aucune donnée structurale n'est disponible pour les autres sous-unités d'eIF3, en particulier les sous-unités fonctionnelles a, b, c, g, i et j.

Les interactions décrites pour ce facteur n'ont pas toutes lieu simultanément. Elles surviennent à différents stades du démarrage mais le moment précis de leur mise en place est encore relativement flou. Un premier complexe stable est vraisemblablement formé entre les eIFs 1, 1A, 2, 3 et 5. Ce complexe appelé MFC (MultiFactor Complex) permet le recrutement de l'ARNt initiateur sur le ribosome (Asano *et al.*, 2000). Un modèle d'interaction a pu être établi pour le MFC de levure à partir des données biochimiques obtenues en utilisant des complexes formés *in vivo* et il est représenté sur la **Figure 1.16** (Vasalek *et al.*, 2003). Les 6 sous-unités fonctionnelles d'eIF3, communes à tous les Eucaryotes, sont positionnées dans ce modèle. Le facteur eIF1A n'est pas indiqué sur ce schéma. Cependant, il interagit très

clairement avec eIF2 et eIF3 et fait donc très probablement partie du MFC (Olsen *et al.*, 2003).



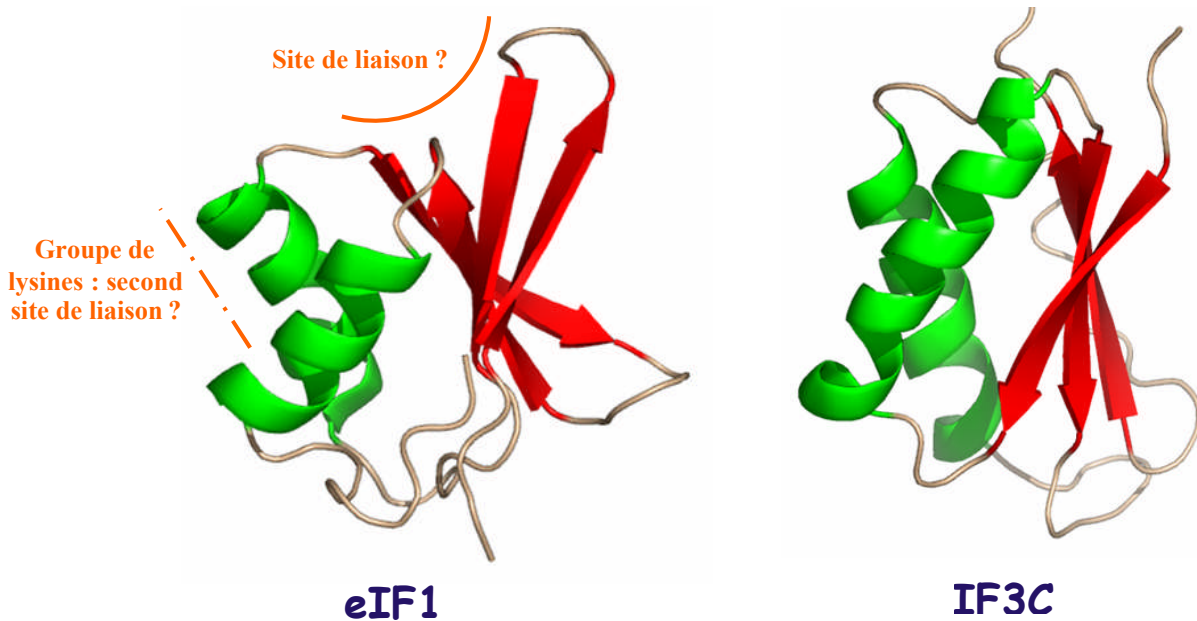
**Figure 1.16 :** Modèle tridimensionnel du MFC d'après Vasalek *et al.*, 2003. Les différentes protéines sont représentées proportionnellement à leur poids moléculaire et le taux de recouvrement entre deux partenaires d'interaction tient compte de la surface d'interaction identifiée biochimiquement.

Ainsi, ce complexe multi-facteurs pourrait agir comme une entité à part entière et se fixer d'un seul bloc à la sous-unité 40S. Néanmoins, les facteurs impliqués dans ce MFC ne se lient probablement pas tous simultanément sur eIF3. De plus, au cours du démarrage, certaines interactions seront rompues sans perturber l'intégrité du complexe.

- **eIF1.**

eIF1 est l'un des plus petits facteurs de démarrage de la traduction eucaryotique (12 kDa). Jusqu'à présent, aucun orthologue du facteur bactérien IF3 n'avait été identifié chez les Eucaryotes et chez les Archées, contrairement aux facteurs IF1 et IF2, universellement conservés. La résolution de la structure du facteur eIF1 humain a permis de reconsidérer ce point de vue (Fletcher *et al.*, 1999). En effet, eIF1 adopte une structure tridimensionnelle très proche de celle du domaine C-terminal d'IF3 (**Figure 1.17**). Bien que ces deux facteurs ne

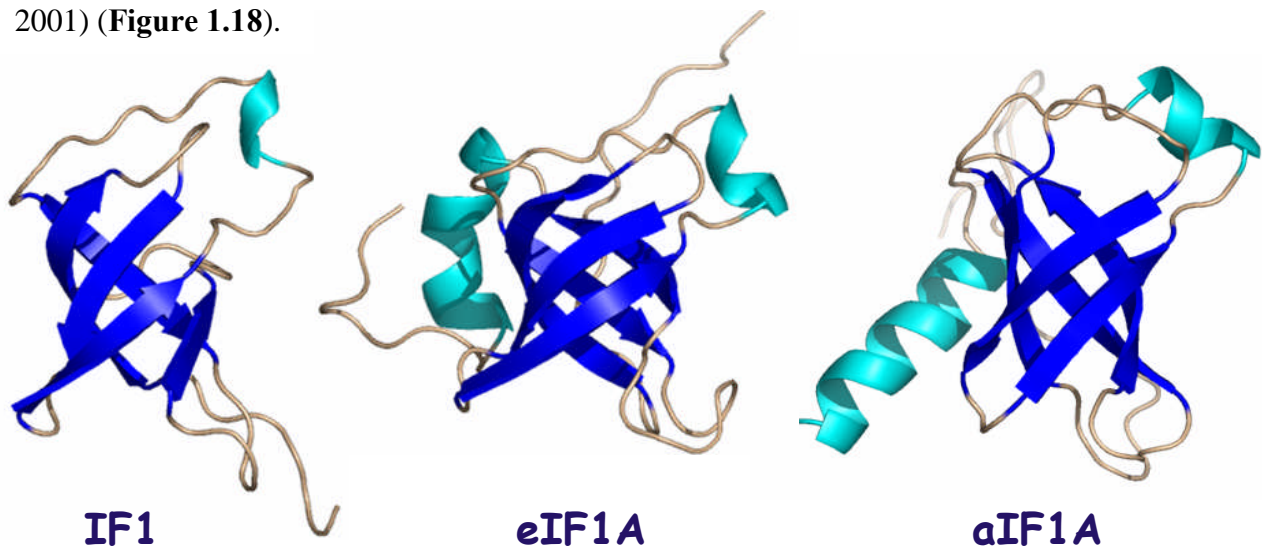
présentent pas d'homologie de séquence, leur ressemblance structurale suggère qu'eIF1 est l'orthologue d'IF3, pour le domaine fonctionnel du facteur bactérien.



**Figure 1.17 :** Structures comparées d'eIF1 humain (Fletcher *et al.*, 1999) et d'IF3C d'*E. coli* (Garcia *et al.*, 1995).

- **eIF1A.**

eIF1A est l'un des deux facteurs de démarrage universellement conservés. Il s'agit d'une protéine de 17 kDa chez les Eucaryotes. Sa structure a été déterminée par RMN pour le facteur eucaryotique et pour son homologue archéen (Battiste *et al.*, 2000 ; Li and Hoffman, 2001) (Figure 1.18).



**Figure 1.18 :** Structure du facteur de démarrage universellement conservé IF1-e/aIF1A. De gauche à droite : structure d'IF1 d'*E. coli* (Sette *et al.*, 1997), structure d'eIF1A humain (Battiste *et al.*, 2000) et structure d'aIF1A de *M. jannaschii* (Li and Hoffman, 2001).

- **eIF2.**

Chez les Eucaryotes, l'ARNt initiateur est spécifiquement reconnu par le facteur de démarrage eIF2. Il s'agit d'un hétérotrimère composé des sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$  (respectivement 36, 38 et 58 kDa chez la levure). Ce facteur est propre à la traduction eucaryotique et archéenne : il n'est pas l'orthologue du facteur IF2 qui assure la fonction de recrutement de l'ARNt initiateur chez les Bactéries. D'ailleurs, leurs mécanismes de reconnaissance de l'ARNt présentent des différences significatives. Tandis qu'IF2 reconnaît l'ARNt<sub>i</sub><sup>Met</sup> sur le ribosome indépendamment de la fixation d'un nucléotide sur son domaine G, eIF2 est capable de sélectionner l'ARNt<sub>i</sub><sup>Met</sup> directement dans le cytoplasme mais cette sélection nécessite la présence impérative d'une molécule de GTP (Kapp and Lorsch, 2003). De plus, le Met-ARNt<sub>i</sub><sup>Met</sup> n'est pas formylé. En conséquence, les déterminants reconnus par eIF2 sur l'ARNt diffèrent de ceux reconnus par IF2. Ainsi, chez les Eucaryotes, eIF2 s'associe à l'ARNt initiateur dans le cytoplasme, sous forme d'un complexe ternaire avec le GTP. Puis, le complexe s'achemine sur le ribosome. Cette fixation sur la 40S est stimulée par les différents facteurs présents dans le MFC (eIF1, eIF1A, eIF3 et eIF5). Il est donc tentant de penser que le complexe ternaire est recruté par le ribosome directement au sein du MFC. En contrepartie, sa présence activerait la fonction anti-associatrice d'eIF3.

Le cas du facteur eIF2 faisant l'objet d'un chapitre complet (**Chapitre II**), il ne sera pas abordé de manière plus approfondie dans cette partie.

## **2. Activité anti-associatrice d'eIF3, eIF1 et eIF1A.**

En fin de synthèse protéique, les ribosomes 80S se retrouvent bloqués sur l'ARNm traduit. Ces ribosomes doivent donc être dissociés en 40S et 60S libérées de leurs composants excédentaires. Comme signalé au début de cette partie, en l'absence d'orthologue eucaryotique de RRF bactérien, le mécanisme par lequel survient cette dissociation reste flou. D'autre part, dans les conditions physiologiques, la concentration en sels présente dans le cytoplasme favorise la réassociation des sous-unités ribosomales. Un mécanisme doit donc empêcher cette réassociation. eIF3, eIF1A et éventuellement eIF1 et eIF6 ont été impliqués dans cette étape.

eIF1 et eIF1A seuls ne possèdent pas de propriétés anti-associatrices. De même, il a été démontré récemment qu'eIF3 seul ne pouvait empêcher l'interaction 40S-60S (Kolupaeva *et al.*, 2005). Cette étude a également montré que l'ajout d'eIF1 et d'eIF1A n'était pas suffisant pour activer la fonction anti-associatrice d'eIF3. En fait, la présence du complexe ternaire eIF2:GTP:Met-ARNt<sub>i</sub><sup>Met</sup> est absolument requise. Ce résultat suggère que c'est la

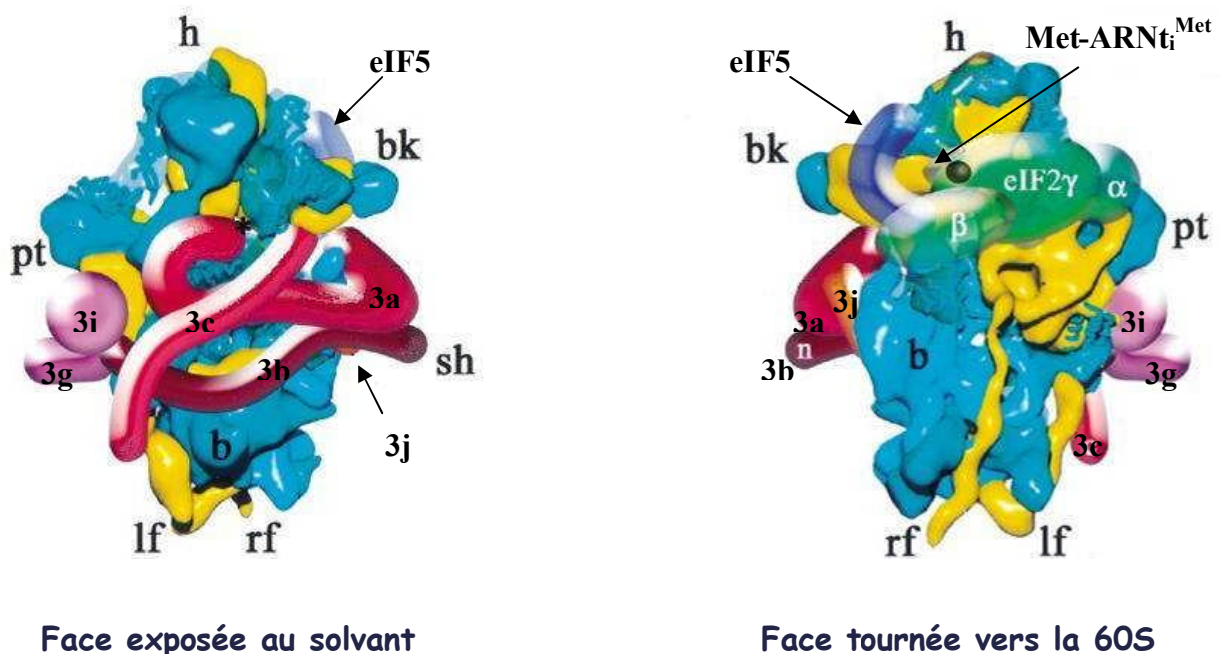
liaison de tout le complexe MFC sur la 40S qui pourrait empêcher sa réassociation avec la 60S. En effet, les interactions entre les différents partenaires du MFC sont très difficiles à scinder les unes des autres. La présence d'un complexe MFC formé de tous les facteurs (eIFs 1, 1A, 2, 3 et 5) est nécessaire pour permettre l'arrimage de l'ARNt-initiateur, et par la suite de l'ARNm, sur le ribosome (Asano *et al.*, 2000).

Du fait de son homologie avec IF1, il a été proposé qu'eIF1A se lie à la 40S de la même manière qu'IF1 sur la 30S. Cette hypothèse placerait donc eIF1A au niveau du site A de la petite sous-unité ribosomale et il pourrait jouer un rôle similaire à celui d'IF1. En particulier, eIF1A pourrait empêcher la dimérisation des sous-unités 40S libérées des complexes 80S, soit par gêne stérique du fait de sa position au niveau de la zone de dimérisation, soit en induisant un changement conformationnel de la 40S qui bloquerait son association avec une autre 40S (Kainuma and Hershey, 2001).

Récemment, eIF1 a également été impliqué dans l'activation de la fonction anti-associatrice d'eIF3 (Kolupaeva *et al.*, 2005). Du fait de la ressemblance structurale entre eIF1 et IF3C, un modèle de positionnement d'eIF1 sur le ribosome a été envisagé par homologie avec la position d'IF3 sur la 30S (Dallas and Noller, 2001). Ce positionnement a ensuite été confirmé par sonde avec des radicaux hydroxyles (Lomakin *et al.*, 2003). Cette étude a ainsi montré qu'eIF1 se place sur la surface de la plateforme de la 40S, au niveau de la zone d'interface avec la 60S, et à proximité du site ribosomal P, soit une position similaire à celle d'IF3C sur la 30S. Ainsi, la présence d'eIF1 à l'interface qui assure le contact avec la 60S pourrait empêcher la grande sous-unité ribosomale d'accéder à des éléments de l'ARNr 18S qui participent à la formation des ponts moléculaires reliant les deux sous-unités.

En se plaçant au niveau de la zone interfacielle de la 40S, eIF1A et eIF1 bloquent donc les interactions 40S-40S et 40S-60S, respectivement. Cependant, ces facteurs seuls n'ont aucune activité anti-associatrice et la présence d'eIF3 est absolument requise pour assurer cette fonction. Ce phénomène peut être dû à un défaut d'association des facteurs à la 40S en l'absence d'eIF3 puisque c'est le complexe MFC qui assure l'arrimage de nombreux facteurs de démarrage sur le ribosome. Cependant, c'est à eIF3 que revient le rôle prépondérant dans cette étape bien que lui-même ne possède pas non plus d'activité anti-associatrice intrinsèque (Chaudhuri *et al.*, 1999). En effet, en l'absence d'eIF1 et d'eIF1A, eIF3 ne peut empêcher la réassociation des 40S et des 60S qu'à condition d'être lié au complexe ternaire eIF2:GTP:Met-ARNt<sub>i</sub><sup>Met</sup> (Kolupaeva *et al.*, 2005). Pour essayer de comprendre la synergie qui existe entre ces 3 facteurs, la position d'eIF3 sur le ribosome a été appréhendée à l'aide de versions du MFC pour lesquelles certaines sous-unités d'eIF3 étaient présentes sous formes tronquées (Valasek *et al.*, 2003). Cette étude a mis en évidence les rôles cruciaux joués par

eIF3a et eIF3c dans la liaison du MFC à la 40S. En particulier, le domaine C-terminal d'eIF3a s'associe spécifiquement avec l'ARNr 18S au niveau des hélices 16 à 18 du polynucléotide. Ces résultats ont conduit à un positionnement du MFC sur la 40S (**Figure 1.19**). Ainsi, eIF3 ne se retrouverait pas sur la face interfaciale de la 40S mais au contraire, il se positionnerait sur la surface de la petite sous-unité ribosomale exposée au solvant. Dans ces conditions, son activité anti-associaatrice ne peut s'exercer par une gêne stérique directe bloquant l'accès de la 40S à la 60S. Plus probablement, eIF3 induit un changement conformationnel au sein de la 40S qui empêche sa réassociation à la 60S et ce mouvement serait stimulé par les facteurs eIF1 et eIF1A ainsi que le complexe ternaire.



**Figure 1.19 :** Modèle d'interaction d'eIF3 avec la 40S d'après Valasek *et al.*, 2003. Les lettres h, bk, sh, rf, lf, b et pt désignent respectivement la tête (head), le bec (beak), l'épaule (shoulder), le pied droit (right foot), le pied gauche (left foot), le corps (body) et la plateforme de la 40S.

Ce positionnement d'eIF3 sur la 40S est cohérent avec le fait qu'eIF3 seul n'est pas suffisant pour empêcher la réassociation des sous-unités ribosomales. D'autres facteurs sont requis pour qu'eIF3 assure cette fonction, et en leur absence, eIF3 pourrait rester associé à la 40S sans gêner la formation des complexes 80S. Plusieurs études ont d'ailleurs suggéré qu'eIF3 reste accroché sur la 40S au cours des premiers cycles de la phase d'élongation et il pourrait également être présent sur les complexes 80S en fin de terminaison de la traduction. Cette stabilité de l'interaction eIF3-40S pourrait s'expliquer par la présence de l'ARNm sur le complexe ribosomal. En effet, très récemment, Kolupaeva *et al.* (2005) ont montré que la liaison d'eIF3 à la 40S était stabilisée par des facteurs oligonucléotidiques d'au moins 25 nucléotides et pouvant éventuellement présenter des structures en épingles à cheveux. Cette

stabilisation a lieu même si la sous-unité eIF3j, qui est pourtant essentielle à une interaction eIF3-40S stable, a été enlevée du facteur (Fraser *et al.*, 2004). En effet, en l'absence de la sous-unité j, le facteur eIF3 ne lie que très faiblement la sous-unité 40S. Toutefois, une interaction efficace est retrouvée pour ce facteur tronqué eIF3j<sup>-</sup> en présence du complexe ternaire eIF2:GTP:Met-ARNt<sub>i</sub><sup>Met</sup> mais également lors de l'ajout d'homopolymères nucléotidiques de type poly-U, poly-dT ou poly-dC. Finalement, au sein des complexes de prédémarrage 48S associés à l'ARNm, eIF3j n'est plus associée au reste du facteur. Cette sous-unité pourrait donc permettre l'arrimage initial du MFC. Puis, lorsque l'ARNm est fixé sur le complexe de démarrage, il stabilise lui-même l'interaction d'eIF3 avec le ribosome et par conséquent, la présence d'eIF3j n'est plus requise.

En l'absence de tout autre facteur de démarrage, ces ARNs homopolymériques sont également capables de stimuler une fonction dissociatrice attribuée à eIF3 (Kolupaeva *et al.*, 2005). Les auteurs ont alors proposé un modèle selon lequel eIF3 jouerait un rôle dans le recyclage du ribosome. En fin de traduction, les complexes 80S sont toujours arrimés sur l'ARNm. La présence de l'ARNm pourrait ainsi stimuler la fixation d'eIF3 sur le ribosome et cette association temporaire activerait la dissociation des 80S par eIF3. Cette hypothèse permettrait d'expliquer la présence d'eIF3 dans les complexes post-terminaux. Dans ce modèle, la dissociation ARNm-stimulée des 80S par eIF3 en cours de traduction serait alors empêchée par la présence d'un ARNt au site P du ribosome. Le facteur eIF3 pourrait d'ailleurs rester associé à la 40S tout au long de ce processus de formation du MFC et de recrutement de l'ARNt initiateur sur la petite sous-unité ribosomale.

### **3. eIF6 : un cas ambigu ?**

eIF6 est un monomère de 26 kDa chez la levure, qui ne possède pas d'homologue bactérien. Cette protéine a longtemps été présentée comme un facteur de démarrage de la traduction. Sa structure a été résolue pour le facteur de *S. cerevisiae* ainsi que pour un homologue archéen (Groft *et al.*, 2000). Ce facteur est effectivement capable de lier les sous-unités 60S et sa présence sur la 60S empêche la formation de 80S actifs dans la traduction (Ceci *et al.*, 2003). Cependant, la déplétion en eIF6 n'entraîne pas un arrêt total de la synthèse protéique comme c'est le cas pour les autres facteurs de démarrage (Si and Maitra, 1999). Dans ces conditions, il ne semble pas être un vrai facteur de démarrage. Une autre étude (Basu *et al.*, 2001) a révélé qu'eIF6 était impliqué dans la biogénèse des sous-unités 60S. Des eIF6 cellulaires sont en effet présents dans le noyau. Le rôle apparent d'eIF6 dans la traduction pourrait donc s'expliquer par son implication dans la synthèse des 60S : en son absence, moins de 60S sont formées et par conséquent, moins de 60S sont disponibles pour la synthèse protéique qui s'en trouve alors ralentie. Toutefois, eIF6 reste associé aux 60S lors de

leur exportation vers le cytoplasme, bloquant leur assemblage avec les 40S. Chez les mammifères, le complexe formé entre le récepteur pour kinases C activées, RACK1, et l'isoforme  $\beta$ II de l'une de ses kinases substrats, PKC, semble conditionner le relargage d'eIF6 de la 60S. Une fois activé, ce complexe RACK1 - PKC  $\beta$ II pourrait phosphoryler la sérine 235 d'eIF6. Le facteur eIF6 ainsi phosphorylé deviendrait inactif et serait éjecté de la 60S, autorisant alors son association avec la 40S (Ceci *et al.*, 2003).

### **3) Recrutement de l'ARNm et balayage de sa séquence.**

Après formation du complexe de prédémarrage 43S et recrutement de l'ARNt initiateur *via* le complexe ternaire eIF2:GTP:Met-ARNt<sub>i</sub><sup>Met</sup>, c'est l'ARNm qui doit s'associer au ribosome. Chez les Bactéries, les phases de transcription et de traduction sont étroitement liées. Ainsi, dès que la séquence SD émerge de la machinerie transcriptionnelle, elle est prise en charge par la petite sous-unité ribosomale qui initie alors la traduction avant même que tout l'ARNm ait été transcrit (Mehdi and Yudkin, 1968). Chez les Eucaryotes, ces deux processus interviennent dans des compartiments cellulaires distincts (noyau pour la transcription et cytoplasme pour la traduction) et ils ne sont pas couplés. Il existe toutefois de rares cas où les processus transcriptionnel et traductionnel peuvent survenir tous les deux au sein du noyau et ils sont alors couplés (Iborra *et al.*, 2004). Dans le cas général, c'est un ARNm entièrement synthétisé, possédant localement des structures secondaires et tertiaires et ayant subi diverses modifications post-transcriptionnelles, qui va entrer dans la machinerie traductionnelle. De plus, cet ARNm ne possède pas de séquence homologue à la séquence SD qui signalerait la position de l'AUG initiateur. Toutefois, chez les mammifères, une séquence en amont de l'AUG initiateur appelée séquence de Kozak est nécessaire à la sélection spécifique du codon de démarrage (Kozak, 1994). Cependant, aucune interaction de type SD-ASD qui faciliterait l'ancrage de l'ARNm sur la 40S n'a pu être mise en évidence pour cette séquence de Kozak. Un tout autre système est donc mis en place pour permettre l'attachement de l'ARNm sur la 40S et la localisation du codon de démarrage. En fait, le ribosome eucaryotique se place à l'extrémité 5' de l'ARNm et va balayer la séquence du polynucléotide dans le sens 5' → 3' jusqu'à rencontrer l'AUG initiateur qui est généralement le premier AUG de la molécule d'ARNm. Cette solution met en jeu une panoplie de facteurs propres au système eucaryotique : les eIF4s.

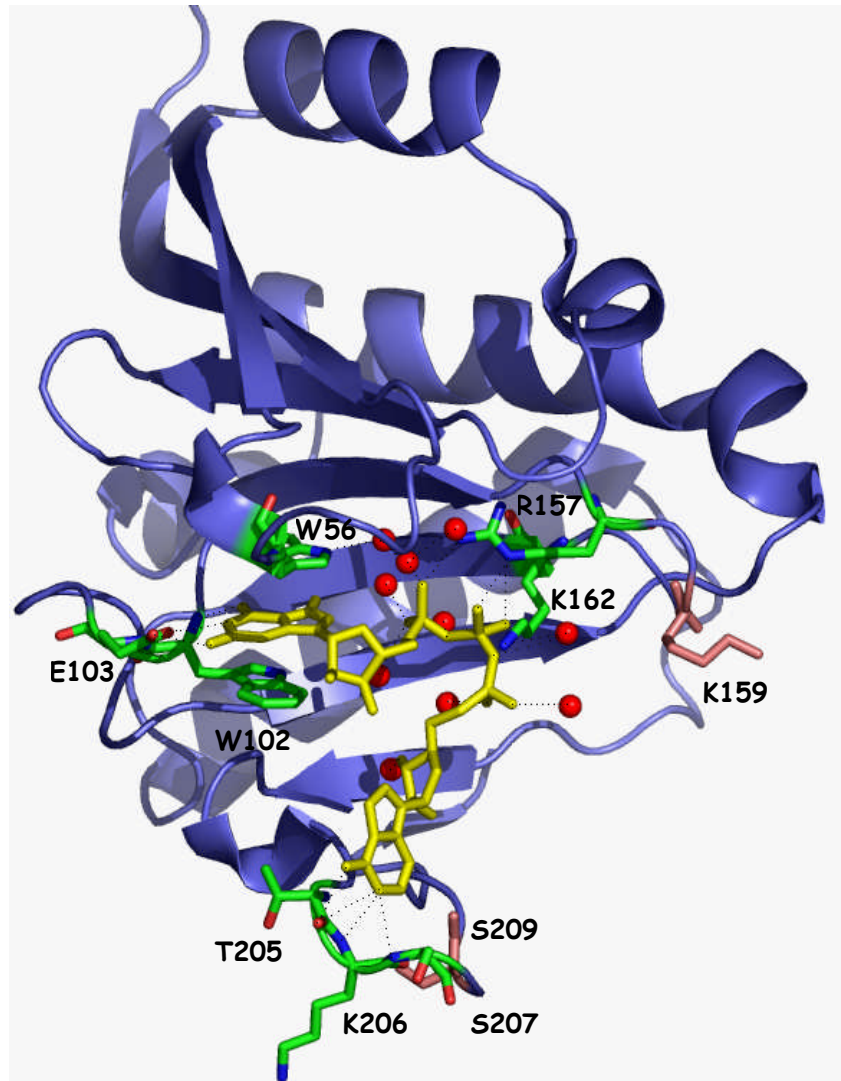
Chez les Eucaryotes, les ARNms transcrits subissent deux modifications post-transcriptionnelles majeures. Premièrement, ils reçoivent une coiffe sur leur extrémité 5' qui correspond à une molécule de guanosine méthylée à la position 7 et qui est connectée au premier nucléotide de l'ARNm par un pont triphosphate 5' à 5' (cette coiffe sera par la suite notée m<sup>7</sup>G(5')ppp(5')N où N désigne le premier nucléotide de l'ARNm). Les ARNms

eucaryotiques n'étant pas directement pris en charge par le ribosome en cours de transcription, cette coiffe est alors essentielle pour empêcher leur attaque par des enzymes de dégradation jusqu'au démarrage de la traduction. La seconde modification concerne l'extrémité 3' de l'ARNm. Après synthèse du transcrit primaire, une endonucléase va tout d'abord venir cliver l'extrémité 3' de l'ARNm. Puis, la poly(A) polymérase, en présence d'ATP, va venir rajouter un nombre variable de nucléotides A en 3' du polynucléotide. Cette queue poly(A) additionnelle atteint généralement 200 à 250 nucléotides chez les mammifères supérieurs. Tout comme la coiffe en 5', elle assure un rôle de protection de l'ARNm vis-à-vis de la dégradation.

### **1. Reconnaissance de la coiffe en 5'.**

La machinerie traductionnelle eucaryotique reconnaît l'extrémité 5' coiffée de l'ARNm. C'est le complexe eIF4F qui se charge de faciliter l'accrochage de l'ARNm sur la 40S. eIF4F est composé de 3 facteurs de démarrage : eIF4A, eIF4E et eIF4G. Au sein du complexe, c'est la sous-unité eIF4E qui est directement responsable de la reconnaissance de la coiffe de l'ARNm. Ce facteur possède une région C-terminale phylogénétiquement conservée d'environ 182 résidus tandis que la longueur et la composition de son extrémité N-terminale sont très variables en fonction des espèces. Plusieurs structures de complexes eIF4E – m<sup>7</sup>G(5')ppp(5')N ont été obtenues au cours des dernières années et en particulier celle du complexe humain (Tomoo *et al.*, 2003) (**Figure 1.20**).

eIF4E ressemble à une main incurvée dans la face concave de laquelle la coiffe vient se fixer. Cette liaison est stabilisée par des interactions de type  $\pi - \pi$  entre la base guanine méthylée en position 7 et les groupes indoles de 2 tryptophanes parfaitement conservés (Trp 56 et 102 dans la **Figure 1.20**). D'autres interactions (liaisons H, contacts de Van der Waals) avec la base G et les 3 groupes phosphates stabilisent l'association de la coiffe sur eIF4E. Cette structure a également montré qu'en présence du premier nucléotide de l'ARNm (mimé ici par la base A), la région C-terminale d'eIF4E (boucle comprenant les résidus 203 à 210) se réorganise et vient interagir avec le nucléotide A. Cette structure permet également de proposer une explication pour l'augmentation d'affinité d'eIF4E vis-à-vis de la coiffe lors de la phosphorylation de sa sérine 209. En effet, cette structure montre que la sérine 209 se trouve à la base de la poche de liaison de la coiffe, au niveau du « tunnel » d'entrée de l'ARNm. Sous forme phosphorylée, elle pourrait interagir avec la lysine 159, de l'autre côté du tunnel (par formation d'un pont salin rétractable ; Marcotrigiano *et al.*, 1997) ou plus vraisemblablement avec un réseau de molécules d'eau, bloquant la sortie du tunnel et empêchant ainsi la coiffe de se dissocier d'eIF4E.

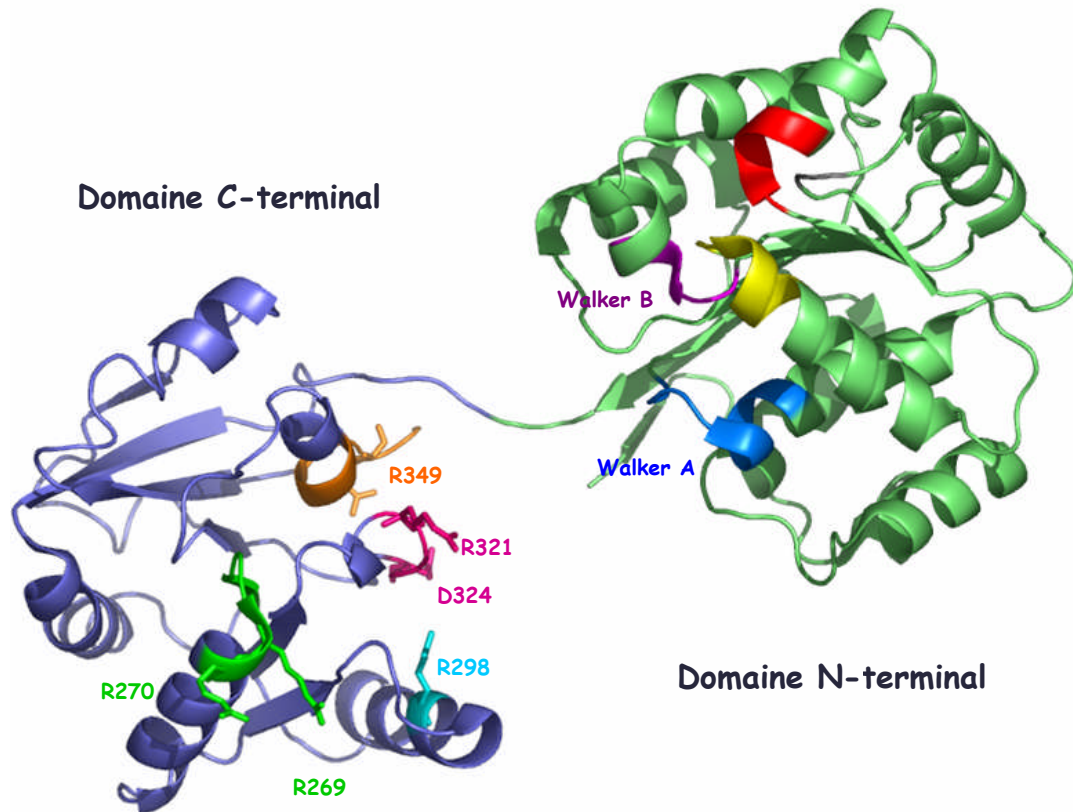


**Figure 1.20 :** Structure du facteur eIF4E humain complexé à un analogue de la coiffe  $m^7GpppA$  à 2.0 Å de résolution (Tomoo *et al.*, 2003). L'analogue de la coiffe est représenté en jaune. Les molécules d'eau et les résidus d'eIF4E qui stabilisent la liaison de la coiffe sont indiqués respectivement en rouge et en vert.

Cette interaction d'eIF4E avec la coiffe est particulièrement importante pour la fixation de l'ARNm sur la 40S mais également pour tout le processus de démarrage car la coiffe stimule la traduction *in vivo* et la force de l'interaction eIF4E -  $m^7GpppN$  pourrait contrôler le taux de synthèse protéique (Niedzwiecka *et al.*, 2002). La liaison d'eIF4E sur l'extrémité 5' coiffée de l'ARNm va permettre de recruter tout le complexe eIF4F sur le polynucléotide. En effet, les ARNm eucaryotiques présentent généralement une région 5' non traduite très structurée. Les duplex ARN-ARN formés dans ces structures secondaires vont gêner la progression de l'ARNm au sein du ribosome. Ces structures secondaires doivent donc être déroulées avant association de l'ARNm à la 40S et c'est l'ARN-hélicase eIF4A qui est responsable de cette fonction.



l'ARN. Sa structure a été résolue pour la protéine de levure (600 résidus) (Caruthers *et al.*, 2000) (**Figure 1.22**).



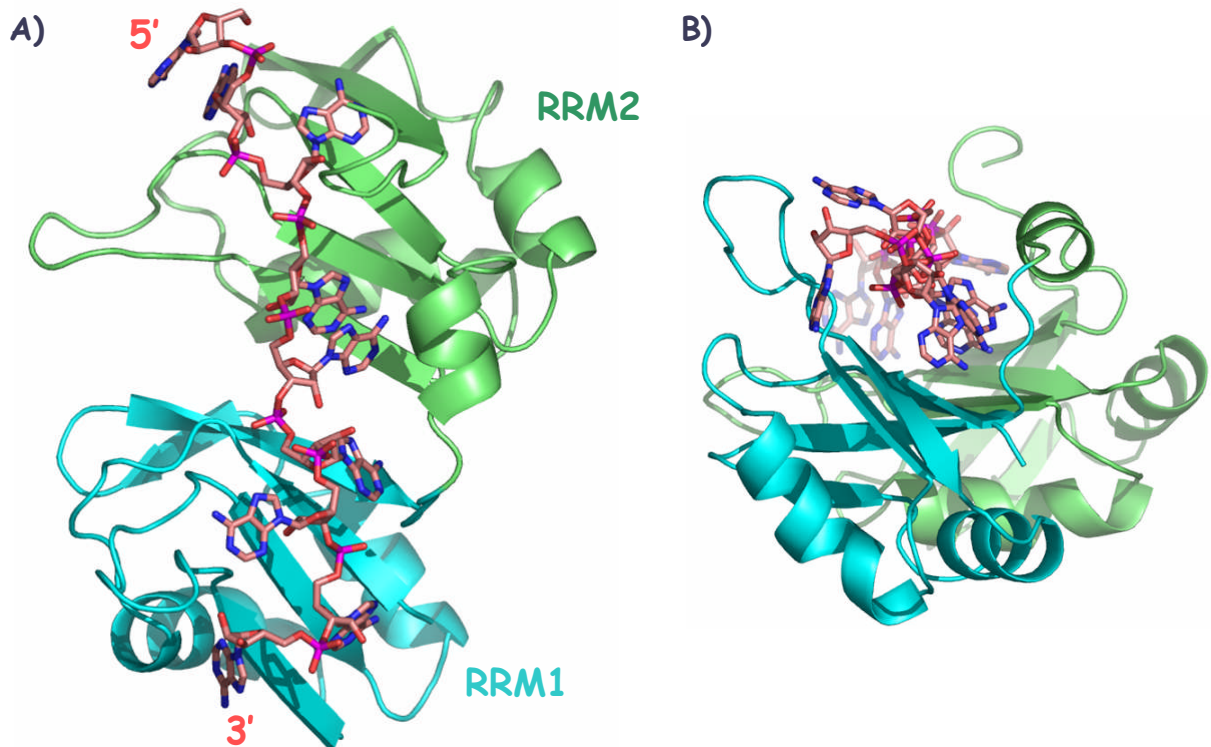
**Figure 1.22 :** Structure d'eIF4A de *Saccharomyces cerevisiae* à 2.8 Å de résolution (Caruthers *et al.*, 2000). Les différents motifs consensus des ARN-hélicases sont colorés comme suit. **Domaine N-terminal :** motif I, Walker A, en bleu ; motif Ia, PTRELA, en jaune ; motif TPGR en rose ; motif II, Walker B (DEAD), en rouge. **Domaine C-terminal :** motif IV, VIFCNTRR, en vert ; motif R-conservé en violet ; motif V, RGID, en magenta ; motif VI, HRIGRGGR, en cyan.

La liaison de l'ATP, son hydrolyse et le relargage d'ADP entraînent divers changements conformationnels au sein d'eIF4A, qui modulent son affinité pour l'ARNm. Ces mouvements pourraient également être utilisés pour transmettre l'énergie libérée lors de l'hydrolyse de l'ATP à la fonction hélicase de la protéine, permettant de rompre les appariements entre les bases de l'ARNm (Lorsch and Herschlag, 1998). En effet, les motifs Walker A et B impliqués dans la liaison et l'hydrolyse de l'ATP se trouvent sur la partie N-terminale de la molécule, tandis que certains motifs impliqués dans la liaison de l'ARNm et qui induisent une activation de la fonction ATPasique d'eIF4A (motif IV, Arg 298) se trouvent dans le domaine C-terminal. Ces deux régions doivent donc communiquer lorsque eIF4A est activé et en particulier, ces changements conformationnels pourraient permettre de rapprocher les deux domaines et de former une protéine compacte autour de l'ARNm.

L'activité ARN-hélicase d'eIF4A seul est cependant peu efficace. *In vitro*, eIF4A ne peut dissocier des duplex de plus de 10 nucléotides et son activité est encore réduite si la séquence du duplex contient un nombre élevé d'appariements forts de type G-C (Rogers *et al.*, 1999). L'incorporation d'eIF4A dans le complexe eIF4F stimule ses 2 activités ATPase et ARN-hélicase. Cependant, l'intervention d'un autre facteur, eIF4B, est nécessaire pour qu'eIF4A ait une efficacité suffisante pour dérouler des duplex ARN-ARN longs et stables. Récemment, une seconde protéine jouant un rôle similaire à eIF4B, eIF4H, a été identifiée chez les mammifères. Ces deux facteurs semblent avoir une fonction redondante dans la traduction. Le modèle actuel propose qu'eIF4A complexé à de l'ATP lie le duplex ARN-ARN à dérouler et, après hydrolyse de l'ATP, rompt quelques appariements entre les premières bases de l'hélice. Puis, eIF4A se déplacerait sur le polynucléotide, soit en utilisant l'énergie de l'hydrolyse de l'ATP, soit par diffusion passive. eIF4B et eIF4H pourraient alors capturer les mono-brins d'ARN libérés par eIF4A et faciliter la dissociation de nouvelles paires de bases. D'autres modèles ont été proposés. En particulier, d'autres ARN-hélicases pourraient être impliquées telles que la protéine Ded1p récemment identifiée comme étant également une ARN-hélicase à boîte DEAD jouant un rôle majeur dans le démarrage de la traduction (Iost *et al.*, 1999).

## **2. Reconnaissance de la queue poly-(A) en 3'.**

En marge des événements à l'extrémité 5' de l'ARNm, la queue poly-(A) en 3' du polynucléotide joue également un rôle crucial pour le recrutement de l'ARNm sur le ribosome. Cette suite de nucléotides A est reconnue par une protéine spécifique, la PABP (Poly-(A) Binding Protein), qui possède 4 domaines RRM (RNA Recognition Motif) se liant à l'ARNm. Cette protéine est d'ailleurs considérée comme un facteur de démarrage à part entière (Kahvejian *et al.*, 2005). Une molécule de PABP se lie à l'ARNm en moyenne tous les 27 nucléotides et les ARNms eucaryotiques possèdent ainsi entre 2 et 10 molécules de PABP sur leur queue poly-(A). La structure de 2 des domaines RRM de la PABP complexés à un polynucléotide de 11 bases A a été déterminée (Deo *et al.*, 1999) (**Figure 1.23**).

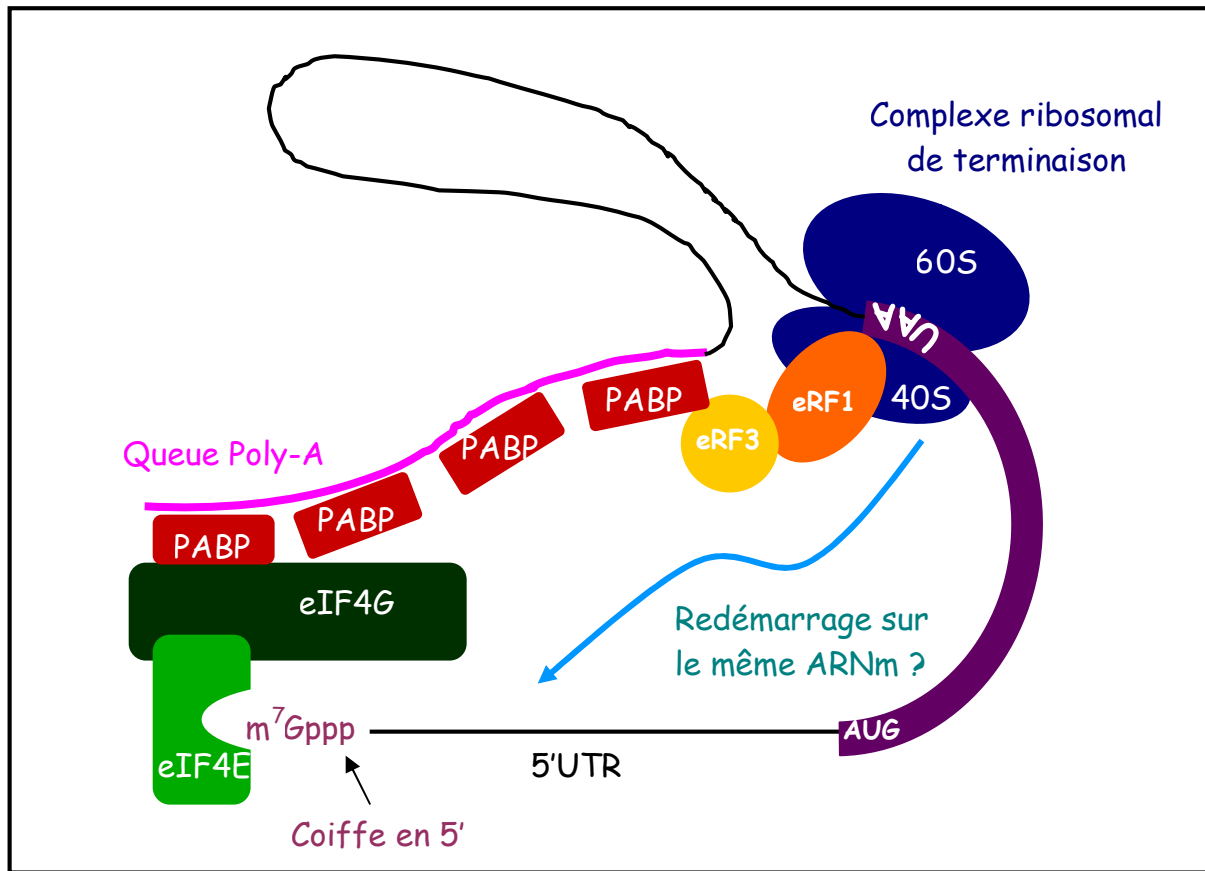


**Figure 1.23 :** Structure des 2 domaines RRM N-terminaux de la protéine humaine de liaison de l'extrémité poly-(A) des ARNm à 2.6 Å de résolution d'après Deo *et al.*, 1999. La vue **B)** correspond à une rotation de **A)** de 90° selon un axe horizontal et de 10° selon un axe vertical.

Au sein de ce complexe, les bases adénines sont reconnues par un jeu de liaisons H, de contacts de Van der Waals et d'interactions aromatiques  $\pi - \pi$  avec des résidus très conservés. Le domaine RRM2 est également reconnu par eIF4G. Cette interaction entre l'extrémité 3' de l'ARNm et une protéine reconnaissant la coiffe en 5' a permis de suggérer un modèle selon lequel l'ARNm serait circularisé au cours de la traduction. La reconnaissance de la queue poly-(A) par la PABP pourrait ainsi servir de signal pour avertir le complexe eIF4F que l'ARNm n'a pas été dégradé et qu'il peut donc être traduit (Wells *et al.*, 1998). D'ailleurs, certains modèles stipulent même qu'en fin de traduction, plutôt que de se dissocier complètement de l'ARNm pour aller traduire un nouveau polynucléotide, le ribosome pourrait se réassocier avec l'extrémité 5' du même ARNm et ainsi, recommencer un cycle de traduction sur le même brin (**Figure 1.24**), ce qui engendrerait un gain d'efficacité considérable pour le processus traductionnel.

Ainsi, la queue poly-(A) et la coiffe de l'ARNm agissent de manière synergétique pour stimuler la liaison du polynucléotide, grâce à leur interaction relayée par l'adaptateur eIF4G. Chez les mammifères, eIF4G assure pour finir un contact direct avec l'autre grand coordinateur du démarrage, eIF3. Cette association, après élimination des structures secondaires de l'ARNm par eIF4A, permet le recrutement définitif de l'ARNm sur le

complexe 43S. Chez les autres Eucaryotes, l'interaction eIF4G-eIF3 pourrait s'effectuer de manière indirecte, par l'intermédiaire d'autres partenaires fixés sur eIF3.



**Figure 1.24 :** Modèle de circularisation de l'ARNm et de redémarrage sur le même ARNm en fin de traduction. eRF1 et eRF3 (pour eukaryotic Release Factor) sont les deux facteurs impliqués dans l'étape de terminaison de la traduction qui entraîne le relargage de la chaîne polypeptidique nouvellement synthétisée.

Une fois ancré sur l'ARNm, le ribosome va pouvoir balayer sa séquence nucléotidique à la recherche de l'AUG. Deux facteurs sont absolument requis pour cette fonction : eIF1 et eIF1A (Pestova *et al.*, 1998). En leur absence, le ribosome est effectivement capable de recruter l'ARNm. Par contre, le complexe de démarrage formé reste bloqué en position 5' de l'ARNm et il est incapable de balayer sa séquence. L'ajout de ces deux facteurs de démarrage permet de dissocier ce complexe inactif pour former un nouveau complexe 48S qui est alors capable de parcourir l'ARNm jusqu'à l'AUG initiateur.

Le mécanisme exact qui permet le balayage de l'ARNm est loin d'être connu. Le seul résultat communément admis sur cette étape est qu'elle nécessite l'hydrolyse d'ATP. Cependant, la fonction précise de cette hydrolyse dans le processus est elle-même méconnue. Ainsi, une enzyme telle qu'eIF4A ou Ded1p pourrait utiliser l'énergie dérivée de l'hydrolyse

de l'ATP pour promouvoir la translocation du ribosome sur l'ARNm. Ou bien, le balayage pourrait se faire par simple diffusion et dans ce cas, l'hydrolyse d'ATP ne servirait qu'à dérouler les structures secondaires de l'ARNm rencontrées en cours de translocation, ces structures se reformant alors après passage du ribosome pour éviter qu'il ne reparte en sens inverse. Ce second modèle serait en accord avec les récentes expériences qui ont montré qu'en présence d'un ARNm très peu structuré sur sa région 5', le codon AUG peut être localisé en l'absence de toute activité ATPasique (sans eIF4A ou Ded1p) (Pestova and Kolupaeva, 2002).

#### **4) Reconnaissance du codon AUG et hydrolyse du GTP.**

Lors du balayage de l'ARNm par le complexe de prédémarrage 48S, le ribosome va finalement stopper sa course sur l'AUG initiateur. Les mécanismes qui permettent au complexe de s'arrêter précisément sur le bon codon de démarrage restent obscurs. Divers composants de la machinerie ont cependant été impliqués dans cette reconnaissance : l'ARNt initiateur bien évidemment, mais aussi les facteurs de démarrage eIF1, eIF2, eIF5 et finalement, les facteurs eIF4G et eIF3 pourraient aussi intervenir. Enfin, chez les mammifères, une séquence en amont de l'AUG appelée séquence de Kozak est également requise.

Chez les mammifères, la reconnaissance du bon codon de démarrage sur l'ARNm nécessite que l'AUG soit dans un contexte nucléotidique favorable (Kozak *et al.*, 1994). Ainsi, la séquence consensus qui permet de signaler un AUG initiateur sur les ARNm est GCC(A/G)CCAAUGG (l'AUG initiateur est souligné) et au sein de cette séquence, les bases indiquées en gras ont un rôle primordial. En effet, lors du processus de balayage, si le premier AUG rencontré n'est pas dans le bon environnement nucléotidique, il peut être contourné par le ribosome qui poursuit alors son chemin en 3'. De nombreuses études ont été menées pour voir si l'ARNr 18S possédait une séquence complémentaire de cette séquence de Kozak qui pourrait alors jouer un rôle identique à la séquence SD des Bactéries. Cependant, ces recherches se sont révélées infructueuses et le mécanisme par laquelle cette séquence influe sur la reconnaissance du codon de démarrage est toujours inconnu. Certains partenaires protéiques tels qu'eIF1 pourraient notamment directement inspecter la séquence nucléotidique entourant l'AUG.

La participation des 3 facteurs de démarrage eIF1, eIF2 et eIF5 à la reconnaissance de l'AUG a tout d'abord été montrée en utilisant un système de criblage (Huang *et al.*, 1997). Dans ce système, on sélectionne des levures devenues capables de démarrer la traduction d'un

gène indicateur sur un codon UUG. Des mutations ont ainsi été isolées dans les gènes codant pour les 3 sous-unités du facteur eIF2 (Donahue *et al.*, 1988 ; Cigan *et al.*, 1989 ; Williams *et al.*, 1989 ; Castilho-Valavicius *et al.*, 1990 ; Dorris *et al.*, 1995) et dans les gènes codant pour eIF1 et eIF5 (Huang *et al.*, 1997). Le cas d'eIF2 sera étudié en détail dans le **Chapitre II** de ce manuscrit.

Les mutations au sein du facteur eIF5 qui sont responsables d'un démarrage sur un codon UUG pourraient provoquer une hyperstimulation de l'activité GTPasique d'eIF2. L'hétérotrimère hydrolyserait ainsi trop tôt sa molécule de GTP et ne réaliserait pas un contrôle efficace de l'appariement codon-anticodon (Huang *et al.*, 1997 ; Unbehaun *et al.*, 2004). De plus, des expériences de « toeprint » ont également suggéré qu'eIF1 et eIF1A étaient nécessaires à la formation de complexes 43S:ARNm capables de localiser le codon AUG (Pestova *et al.*, 1998). La réalisation de ces expériences de toeprint en l'absence d'eIF1 conduit à l'assemblage de complexes 43S à la fois sur des codons AUG et sur d'autres types de codons (Pestova and Kolupaeva, 2002).

Dans le cas du facteur eIF1, les mutations permettant le démarrage sur un codon UUG correspondent toutes à une même zone de la protéine où se concentrent un grand nombre de résidus très conservés (premier site de liaison indiqué en orange sur la **Figure 1.17**). Cette région pourrait ainsi constituer un site de liaison pour un autre partenaire de la machinerie traductionnelle. eIF1 possède également une série de lysines le long de la première hélice de sa structure et ces lysines pourraient très facilement interagir avec le squelette phosphaté d'un ARN (Fletcher *et al.*, 1999). Il pourrait donc s'agir de l'ARNr mais aussi de l'ARNm ou même de l'ARNt. Ainsi, eIF1 pourrait coupler l'appariement codon-anticodon à l'hydrolyse du GTP lié à eIF2 (Unbehaun *et al.*, 2004). eIF1 pourrait également interagir avec la tige anticodon de l'ARNt ou une autre région du polynucléotide et détecter un changement conformationnel de ce dernier lors de l'appariement avec l'AUG. Enfin, le facteur pourrait s'associer directement avec l'ARNm et éventuellement reconnaître la séquence de Kozak. Cependant, aucune liaison à un oligonucléotide portant la séquence GCCACAAUGGCA n'a été détectée *in vitro* (Fletcher *et al.*, 1999).

Plus récemment, des expériences ont suggéré qu'eIF1 se localise près du site P de la 40S (Lomakin *et al.*, 2003). Comme nous l'avons déjà signalé, eIF1 présente une organisation structurale très proche de celle du domaine C-terminal d'IF3 bactérien. Comme IF3C, il pourrait directement inspecter l'appariement codon-anticodon. En présence de 3 paires de bases parfaitement appariées, il entraînerait alors l'arrêt du balayage et informerait eIF2 et/ou eIF5 que l'hydrolyse du GTP peut avoir lieu.

La transmission d'information entre eIF1 et eIF2 est mal comprise. Une interaction directe entre eIF1 et eIF5 reste très controversée. Cependant, le modèle de positionnement de tous ces facteurs au sein du MFC montre qu'eIF1 se retrouve à proximité du domaine C-terminal d'eIF5 lorsque les deux facteurs sont liés à eIF3 (**Figure 1.16**). Du fait de ce rapprochement, le relais de l'information semble facilité. Le travail de Unbehaun *et al.* (2004) montre qu'eIF1 jouerait en réalité le rôle d'inhibiteur de l'hydrolyse du GTP, probablement en influençant la GAP eIF5. Il empêcherait ainsi une hydrolyse prématurée du nucléotide guanine porté par eIF2 et cette inhibition serait levée lorsque l'appariement correct codon-anticodon est réalisé. Une étude thermodynamique des mécanismes impliqués dans l'hydrolyse du GTP, à partir d'un système simplifié, complète l'étude du rôle d'eIF1 (Maag *et al.*, 2005). En effet, lors de la reconnaissance de l'AUG, un changement conformationnel survient au sein du complexe 43S-ARNm. Ce changement induirait une baisse significative de l'affinité d'eIF1 pour le ribosome et dans un second temps, ce facteur serait relargué de la 40S. Son éjection lèverait alors l'inhibition exercée sur l'hydrolyse du GTP. En cas d'appariement avec un codon non-AUG, l'affinité d'eIF1 pour la 40S ne serait pas suffisamment réduite pour entraîner son éjection et l'inhibition de l'hydrolyse du GTP ne serait par conséquent pas levée. Le moment exact auquel survient le relargage d'eIF1 du ribosome reste toutefois incertain (Unbehaun *et al.*, 2004 ; Maag *et al.*, 2005).

Ainsi, eIF1 joue un rôle essentiel dans la sélection de l'AUG initiateur et dans la coordination de cette reconnaissance avec l'hydrolyse du GTP. eIF3 et eIF4G pourraient aussi participer à cette étape (He *et al.*, 2003). Cependant, du fait de leur rôle général de coordinateur, cette implication pourrait être indirecte et les effets observés pourraient résulter d'une simple stabilisation des complexes en présence de ces deux facteurs. Après hydrolyse du GTP stimulée par eIF5, l'affinité du complexe binaire eIF2-GDP pour l'ARNt est fortement réduite et le facteur est éjecté de la 40S en même temps qu'eIF5. Le devenir des 2 autres facteurs de démarrage eIF1 et eIF3 est plus flou. eIF1A reste accroché sur la 40S et va interagir avec eIF5B qui permet l'association des sous-unités ribosomales.

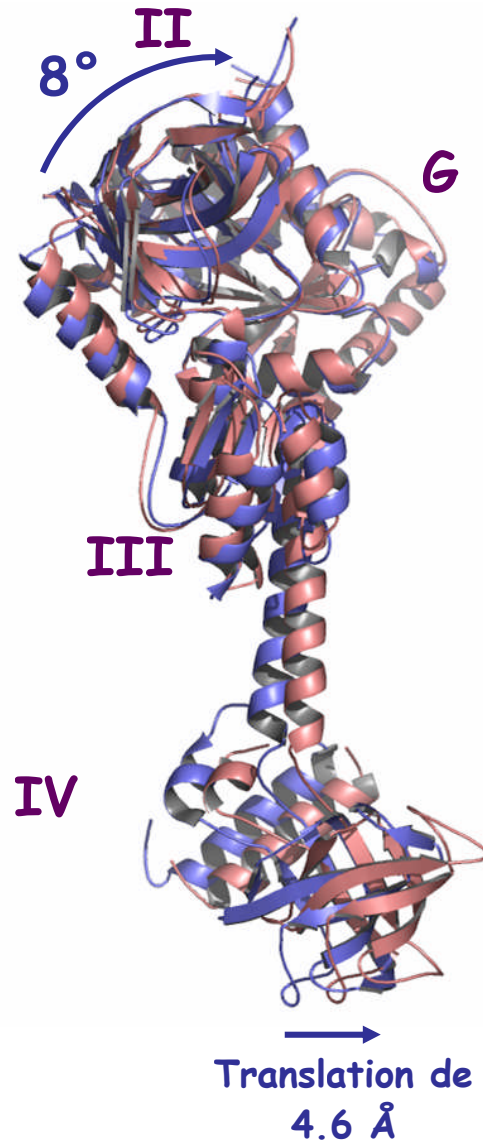
### **5) Association des sous-unités ribosomales.**

La communauté scientifique travaillant sur le démarrage de la traduction a longtemps considéré qu'une seule réaction d'hydrolyse du GTP survenait au cours de cette étape. Ainsi, l'hydrolyse du GTP porté par eIF2 aurait à la fois permis le contrôle de la sélection du bon AUG, le relargage des facteurs de démarrage et l'assemblage des sous-unités ribosomales. Cependant, la découverte d'un orthologue eucaryotique du facteur bactérien IF2 a remis en cause cette vision. En effet, eIF5B intervient dans la dernière étape du démarrage : la formation du complexe 80S (Pestova *et al.*, 2000). C'est un facteur de démarrage à part

entière et également une GTPase (Lee *et al.*, 1999), au même titre qu'IF2. Il a ainsi été démontré, à l'aide d'un système de traduction mammifère *in vitro*, que deux molécules de GTP sont hydrolysées au cours du processus de démarrage (Lorsch and Herschlag, 1999) et c'est eIF5B qui catalyse la seconde réaction GTP-dépendante (Lee *et al.*, 2002).

Du fait de son homologie avec IF2, eIF5B est supposé fonctionner selon un mode similaire à celui du facteur bactérien lors de l'association des sous-unités ribosomales. La structure du facteur a été résolue pour la protéine archéenne aIF5B de *M. thermoautotrophicum* (Roll-Mecak *et al.*, 2000) (voir **Figure 1.7**). Cette structure a permis de souligner les changements conformationnels qui surviennent lors du passage de la forme GTP à la forme GDP du facteur (voir discussion du paragraphe 5 de la Partie A de ce chapitre et **Figure 1.25**). Cependant, l'hydrolyse du GTP n'est pas directement liée à la phase d'assemblage des 40S et des 60S. En effet, des mutations au sein du domaine G d'eIF5B entraînent un sévère défaut d'activité GTPasique mais ne bloquent pas l'association des sous-unités ribosomales (Shin *et al.*, 2002). En réalité, comme pour le facteur IF2 bactérien, l'hydrolyse du GTP ne survient qu'après formation des 80S. Les mouvements structuraux au sein d'eIF5B qui surviennent après hydrolyse du GTP permettent de moduler son affinité pour le ribosome : sous forme GDP, l'affinité d'eIF5B pour le complexe 80S est réduite, ce qui permet son éjection et celle des autres facteurs de démarrage encore présents sur le ribosome (eIF1A, et éventuellement eIF1 et eIF3). Cependant, au sein du complexe bactérien 70S obtenu par cryo-EM (Allen *et al.*, 2005), IF2 adopte une conformation différente de celle observée pour son orthologue aIF5B en solution. Ainsi, les réarrangements conformationnels d'aIF5B qui surviennent entre les états GTP et GDP du facteur ne reflètent probablement pas le mouvement qui a réellement lieu sur le ribosome. Par contre, ils témoignent de la propension qu'a l'hydrolyse du GTP à faire bouger le domaine C-terminal du facteur, ce qui expliquerait les effets observés sur le ribosome.

Il a également été proposé qu'eIF5B facilite l'ajustement de l'ARNt initiateur au site P du ribosome (Schmitt *et al.*, 2002). Or, tandis qu'IF2 bactérien est capable de lier l'ARNt initiateur en solution et sur le ribosome, aucune interaction directe entre l'ARNt et eIF5B n'avait été mise en évidence jusqu'à présent. Toutefois, une étude récente réalisée au laboratoire montre qu'eIF5B est effectivement capable de lier l'extrémité 3' méthionylée de l'ARNt initiateur et que cette liaison, comme dans le cas d'IF2 bactérien, s'établit grâce au domaine IV (VI-2 chez IF2) de la protéine (Guillon *et al.*, 2005). Ainsi, eIF5B pourrait interagir avec l'ARNt initiateur pour le positionner précisément au site ribosomal P lors de l'ancrage de la 60S sur la petite sous-unité ribosomale.



**Figure 1.25 :** Mouvements conformationnels induits lors de l'hydrolyse du GTP par eIF5B. En violet, structure du complexe aIF5B-GDP de *M. thermoautotrophicum*. En rose, structure du complexe aIF5B-GTP de *M. thermoautotrophicum* (Roll-Mecak *et al.*, 2000).

Similairement à IF2 qui interagit avec IF1, eIF5B interagit avec eIF1A. Chez les Eucaryotes, une interaction supplémentaire à celle observée entre les facteurs bactériens a lieu (Marintchev *et al.*, 2003). En effet, le facteur IF1 lie le domaine V d'IF2 (qui correspond au domaine II d'e/aIF5B) *via* la zone concave de son OB-fold. Cette interaction existe également entre eIF1A et eIF5B. Cependant, le domaine C-terminal d'eIF1A (absent chez IF1) interagit aussi avec les hélices additionnelles du domaine C-terminal d'eIF5B (absentes chez IF2). Cette seconde interaction est donc propre aux Eucaryotes et pourrait avoir une implication fonctionnelle significative par rapport au mécanisme bactérien. De plus, eIF1A se trouve au site A de la 40S et son interaction avec eIF5B permet de positionner eIF5B près du site P du ribosome, ce qui serait cohérent avec un rôle du facteur dans l'ajustement de la position de

l'ARNt initiateur sur le ribosome. Dans tous les cas, ces deux facteurs agissent en étroite collaboration lors de l'étape d'assemblage du complexe 80S.

De nombreuses questions restent encore posées sur le déroulement mécanistique de cette étape. En effet, le moment exact auquel eIF5B est recruté sur le ribosome est encore inconnu : l'hydrolyse du GTP porté par eIF2 et le relargage du facteur sont-ils un élément déclenchant ? Quoi qu'il en soit, l'activité GTPasique d'eIF5B est étroitement couplée à la fidélité du processus de démarrage. Un relargage prématuré du facteur de la 80S dû à un défaut de son activité GTPasique provoque la reprise du balayage de l'ARNm par le ribosome : l'AUG initiateur est contourné et le ribosome poursuit sa recherche en aval (Shin et al., 2002). Ceci pourrait être la conséquence de l'absence d'ajustement de l'ARNt initiateur au site P. Dans ce cas, la reconnaissance de l'AUG initiateur ne serait pas achevée et le ribosome poursuivrait son chemin sur le polynucléotide.

D'autres études ont également mis en cause la PABP dans cette étape (Proweller and Butler, 1996). En effet, la PABP pourrait se fixer aux sous-unités 60S indépendamment de la queue poly-(A) et ainsi favoriser l'association des sous-unités ribosomales.

Ainsi, de nombreux points restent à éclaircir concernant le mode d'action exact d'eIF5B. Toutefois, après hydrolyse de sa molécule de GTP et relargage des derniers facteurs de démarrage, un complexe 80S assemblé sur le bon codon AUG est libéré et il va pouvoir débiter la phase d'élongation.

La dernière étape à accomplir pour clore ce démarrage concerne le recyclage d'eIF2. En effet, eIF2 est relargué de la 40S sous forme d'un complexe binaire avec une molécule de GDP. Cette forme d'eIF2 présente une affinité réduite pour l'ARNt initiateur par rapport à la forme « active » eIF2-GTP. Dans les conditions physiologiques, le recyclage d'eIF2-GDP en eIF2-GTP aurait lieu trop lentement. Un facteur est donc nécessaire pour accélérer le processus : le facteur d'échange eIF2B. Ce facteur catalyse l'échange  $\text{GDP} \rightarrow \text{GTP}$  sur eIF2 et sa fonction joue un rôle primordial dans la régulation de tout le processus de traduction (Williams *et al.*, 2001). Cette étape sera abordée plus en détails dans le **Chapitre II**.

Enfin, il faut noter que certaines autres protéines ont été potentiellement impliquées dans le démarrage de la traduction, en particulier eIF2A, eIF2C et eIF5A. eIF2A et eIF2C pourraient recruter l'ARNt initiateur tout comme eIF2, mais ces facteurs interviendraient dans un mode de démarrage parallèle au mode canonique (Zoll *et al.*, 2002 ; Bagchi *et al.*, 2005). eIF5A pourrait participer à la formation de la première liaison peptidique (Kang and Hershey, 1994). Enfin, eIF2C, tout comme eIF6, serait également impliqué dans la biogénèse des 60S.

Cependant, aucun résultat expérimental n'a permis d'affirmer jusqu'à présent que ces protéines assument un véritable rôle de facteur de démarrage. C'est pourquoi leur fonction détaillée ne sera pas abordée ici.

## **6) Autres modes de démarrage.**

Le mécanisme de démarrage de la traduction eucaryotique décrit jusqu'à présent est le mécanisme le plus fréquent. Cependant, d'autres mécanismes, minoritaires, existent.

A l'heure actuelle, le processus le plus répandu parmi ces modes alternatifs est appelé démarrage à partir d'un site interne d'entrée du ribosome (IRES pour Internal Ribosomal Entry Site). Dans ce mécanisme indépendant de la coiffe en 5' de l'ARNm, le balayage classique de l'ARNm n'a pas lieu et bien souvent, un panel réduit d'eIFs est suffisant (Pisarev *et al.*, 2005 pour revue). Les IRES correspondent à de longues portions d'ARN qui présentent une structure secondaire très étendue qui va permettre l'accrochage direct de l'ARNm au ribosome. Ces IRES se situent généralement en amont du second cistron d'un ARNm bicistronique et ils permettent la traduction de ce second cistron indépendamment de celle du premier cistron, qui suit généralement un mode canonique, coiffe-dépendant. Ce mode de démarrage est fréquemment rencontré pour la traduction d'ARNms viraux tels que les ARNms du virus de l'Hépatite C (HCV) et du virus de la paralysie du cricket (CrPV).

Il existe également un mode de démarrage eucaryotique appelé « shunting » utilisé pour démarrer la traduction de divers ARNms viraux polycistroniques et très structurés. Dans ce processus, le complexe de prédémarrage se fixe au niveau de la coiffe en 5' de l'ARNm. Cependant, le balayage de l'ARNm ne s'effectue pas de manière linéaire et le complexe ribosomal est en fait capable de contourner les différentes régions très structurées et inhibitrices pour aller se fixer sur un codon AUG initiateur en aval du polynucléotide (Futterer *et al.*, 1993).

Enfin, le « leaky scanning » est un mode de démarrage rencontré sur des ARNms polycistroniques pour lesquels le premier AUG initiateur est dans un contexte nucléotidique faible. Dans ce cas, soit le complexe ribosomal le contourne pour aller se fixer sur un AUG initiateur plus en aval, soit la première région codante est traduite mais après traduction, la sous-unité 40S reste fixée sur le ribosome et un redémarrage de la traduction a alors lieu sur l'un des codons initiateurs en aval (Kozak, 2002).

Ces différents mécanismes ne seront pas décrits en détails ici car l'étude du facteur e/aIF2 menée au cours de ce travail de thèse s'inscrit dans le cadre d'un démarrage canonique

de la traduction, c'est-à-dire 5'-dépendant dans le cas des Eucaryotes. Ils permettent cependant de souligner la complexité de cette étape de démarrage chez les Eucaryotes et la mise en place de divers systèmes pour assurer une synthèse protéique la plus fidèle et la plus efficace possible.

## **Partie C : Le démarrage de la traduction chez les Archées.**

Le monde vivant a longtemps été divisé en deux grandes catégories : les Bactéries et les Eucaryotes. Cependant, la découverte des Archées à la fin des années 70 a conduit le monde scientifique à remettre en cause ce modèle bipartite (Woese and Fox, 1977). Ces organismes, qui vivent souvent dans des conditions extrêmes (le domaine archéen regroupe les méthanogènes, les halophiles et les thermoacidophiles), ont d'abord été classifiés au sein des Bactéries. En réalité, ils possèdent des caractéristiques eucaryotiques, en plus de leurs attributs bactériens. Ainsi, le séquençage complet des premiers génomes d'Archées a révélé la présence de gènes codant pour des protéines homologues à des protéines eucaryotiques et bactériennes. Du point de vue de l'évolution, ces organismes devaient donc être classés dans une troisième lignée de descendance depuis l'ancêtre commun, distincte des lignées eucaryotique et bactérienne (Woese *et al.*, 1978 ; Woese *et al.*, 1990).

### **1) La traduction chez les Archées : un mécanisme hybride.**

Le processus traductionnel des Archées et notamment, l'étape de démarrage de la synthèse protéique, présentent des caractéristiques à la fois bactériennes et eucaryotiques. En effet, les ARNm archéens peuvent être polycistroniques et ils présentent généralement une séquence de type Shine-Dalgarno en amont de l'AUG initiateur. Cependant, les séquençages de génomes d'Archées ont montré que les facteurs de démarrage mis en jeu au cours de cette étape étaient essentiellement homologues aux facteurs eucaryotiques (à l'exception des facteurs universellement conservés). Le **Tableau 1.26** répertorie les différents facteurs de démarrage de la traduction dans les trois domaines du monde vivant, en mettant en parallèle les homologies entre domaines.

Ainsi, l'étude du mécanisme de démarrage archéen se révèle particulièrement intéressante sur deux points : d'une part, elle permet d'apporter de nouvelles informations sur les mécanismes à la fois eucaryotique et bactérien ; d'autre part, la compréhension des adaptations qui ont été réalisées au sein de ces organismes face à leurs conditions d'évolution permettra d'apporter un nouveau regard sur le mécanisme de la traduction au sein de l'ancêtre commun LUCA (Last Universal Common Ancestor) (Londei, 2005 ; Kapp and Lorsch, 2004 ; pour revues).

Notation bactérienne	BACTERIA	Notation eucaryotique	EUKARYA	ARCHAEA
	Domaine C-terminal d'IF3	Initiation Factor 1	eIF1	aIF1
Initiation Factor 1	IF1	Initiation Factor 1A	eIF1A	aIF1A
Initiation Factor 2	IF2	Initiation Factor 5B	eIF5B	aIF5B
		Initiation Factor 2	eIF2 ( $\alpha, \beta, \gamma$ )	aIF2 ( $\alpha, \beta, \gamma$ )
		Initiation Factor 2A	eIF2A	
		Initiation Factor 2B	eIF2B ( $\alpha, \beta, \gamma, \delta, \epsilon$ )	aIF2B ( $\alpha, \beta, \delta$ )
		Initiation Factor 2C	eIF2C	
Initiation Factor 3	IF3			
		Initiation Factor 3	eIF3	
		Initiation Factor 4A	eIF4A	
		Initiation Factor 4B	eIF4B	
		Initiation Factor 4E	eIF4E	
		Initiation Factor 4G	eIF4G	
		Initiation Factor 4H	eIF4H	
		Initiation Factor 5	eIF5	
Elongation Factor P	EF-P	Initiation Factor 5A	eIF5A	aIF5A
		Initiation Factor 6	eIF6	aIF6
		Poly-(A) Binding Protein	PAPB	
			Ded1p	

**Tableau 1.26 :** Facteurs de démarrage de la traduction dans les trois domaines du monde vivant.

## **2) Démarrage de la traduction archéenne : similitudes et divergences avec les autres domaines.**

A ce jour, plus d'une vingtaine de génomes d'Archées ont été entièrement séquencés. Mais peu de données expérimentales sont disponibles sur la traduction archéenne, en partie car la génétique moléculaire est encore peu développée chez ces organismes. Ainsi, beaucoup de schémas sur le mécanisme général du processus ont été tracés par comparaison avec les deux autres systèmes détaillés dans les parties A et B de ce chapitre. Cependant, un intérêt

croissant est aujourd'hui porté aux Archées et des données biochimiques et structurales ont émergé au cours des dernières années, permettant de dresser un premier modèle pour ce mécanisme singulier.

### **1. L'ARNm.**

Les Archées ne possèdent pas de noyau. En conséquence, comme chez les Bactéries, un couplage entre la transcription et la traduction a probablement lieu. Les ARNm archéens ne possèdent pas de coiffe en 5'. Ils présentent une organisation polycistronique qui rappelle celle des ARNm bactériens et une séquence de type Shine-Dalgarno est souvent rencontrée en amont du codon initiateur (Keeling and Doolittle, 1995). En conséquence, un mode de reconnaissance du codon de démarrage similaire à celui mis en place chez les Bactéries est envisagé. Ce mécanisme passerait par une reconnaissance directe de la position de l'AUG *via* l'appariement entre la séquence SD de l'ARNm et la région complémentaire de l'ARN ribosomal 16S (Condo *et al.*, 1999) D'ailleurs, contrairement au cas bactérien, l'ARNm archéen peut former un complexe binaire stable avec la 30S en l'absence de tout autre composant traductionnel (Benelli *et al.*, 2003).

Cependant, les Archées possèdent également de nombreux ARNm pour lesquels presque toute l'extrémité 5' non traduite a été tronquée. Ces ARNm sans 5'-UTR (leaderless mRNAs) sont également présents dans les autres domaines du monde vivant. Toutefois, leur proportion par rapport aux ARNm canoniques est particulièrement élevée chez les Archées. Pour ces ARNm, la présence du seul ARNt initiateur permet leur interaction avec la petite sous-unité ribosomale (Benelli *et al.*, 2003). De plus, des ARNm possédant une région 5' entière mais pour lesquels la séquence SD a été enlevée ne sont plus traduits par les cellules archéennes. Ils peuvent cependant réintégrer le processus traductionnel à condition que leur séquence 5'-UTR soit entièrement supprimée (Condo *et al.*, 1999). Ceci suggère qu'un autre mode de démarrage existe au sein de ces cellules pour traduire les ARNm sans région 5'UTR. D'ailleurs, ce mode de démarrage pourrait être un vestige du mécanisme de l'ancêtre commun LUCA puisque l'on trouve de tels ARNm dans les trois domaines du monde vivant. Pour une raison encore inconnue, ce mécanisme primitif serait encore relativement développé chez les Archées comparativement aux autres espèces (Londei, 2005).

### **2. Les facteurs de démarrage.**

Tout comme les Bactéries et les Eucaryotes, les Archées possèdent les facteurs de démarrage de la traduction universellement conservés aIF1A et aIF5B. Un homologue d'eIF1 est également présent. Le facteur aIF5A qui est homologue à eIF5A et au facteur d'élongation

bactérien EF-P intervient aussi chez les Archées. Les autres facteurs de démarrage rencontrés chez ces organismes sont exclusivement homologues aux facteurs eucaryotiques : le facteur hétérotrimérique e/aIF2, le facteur anti-associateur e/aIF6 et les trois sous-unités régulatrices du facteur d'échange e/aIF2B. Comparativement aux Eucaryotes, les facteurs manquants sont le multicomplexe eIF3, la GAP eIF5, les protéines de la coiffe eIF4F, eIF4B, eIF4H et la PABP, et les sous-unités catalytiques  $\gamma$  et  $\epsilon$  du facteur d'échange eIF2B.

Ainsi, en dehors des facteurs universellement conservés, la traduction archéenne fait appel à l'hétérotrimère aIF2 ainsi qu'au facteur aIF1, malgré un mode de reconnaissance de l'ARNm probablement similaire au mode bactérien. Comme nous l'avons précédemment souligné, eIF2 permet la prise en charge spécifique de l'ARNt initiateur dans les cellules eucaryotiques et eIF1 est impliqué dans le processus de balayage de l'ARNm ainsi que dans la reconnaissance du codon AUG potentiellement *via* une inspection directe de l'appariement codon-anticodon. Ainsi, la présence d'homologues de ces facteurs chez les Archées pourrait traduire le fait que l'ARNt initiateur archéen est plus proche de l'ARNt initiateur eucaryotique que de l'ARNt<sub>f</sub><sup>Met</sup>. En effet, chez les Archées, la méthionine portée par l'ARNt initiateur n'est pas formylée (Ramesh and RajBhandary, 2001). De plus, cet ARNt présente la paire de bases A1:U72 universellement conservée chez tous les ARNts initiateurs eucaryotiques. L'absence de formylation et les caractéristiques proches de celles du polynucléotide eucaryotique sont corrélées à la prise en charge de l'ARNt initiateur archéen par un facteur hétérotrimérique aIF2 homologue au facteur eIF2.

A l'inverse, deux facteurs typiquement eucaryotiques sont absents chez les Archées : eIF5 et eIF3. L'absence d'eIF5 est particulièrement surprenante puisque les Archées possèdent un homologue d'eIF2 (Dever, 2002). Chez les Eucaryotes, eIF5 est absolument nécessaire pour activer la fonction GTPasique d'eIF2. Son omission dans le mécanisme archéen suggère soit qu'aIF2 seul est capable d'hydrolyser le GTP, soit qu'il existe un autre facteur jouant le rôle de GAP chez les Archées, mais celui-ci ne serait alors pas homologue à eIF5. Enfin, comme dans le cas eucaryotique, la sous-unité ribosomale 30S pourrait également intervenir en tant que GAP dans le mécanisme d'hydrolyse du GTP par aIF2.

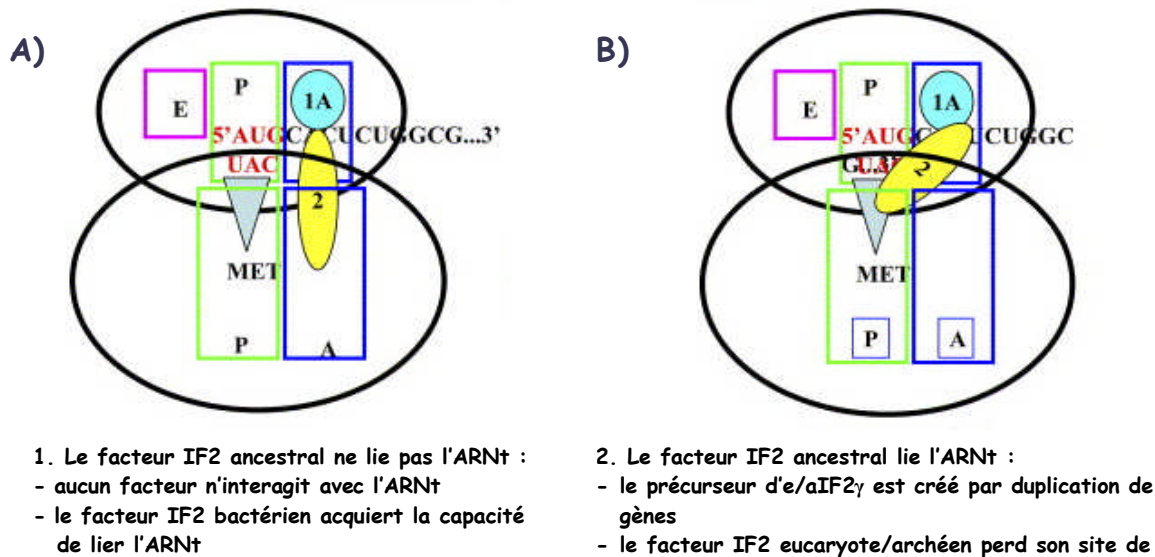
L'absence d'un homologue d'eIF3 (ou tout au moins des sous-unités centrales du facteur) est surprenante. En effet, ce multimère permet de relier tous les partenaires du démarrage de la traduction entre eux et stabilise leur arrimage sur le ribosome. Le système archéen semble pouvoir se passer d'un tel adaptateur. Il faut cependant noter que le facteur eIF3 joue également un rôle majeur dans l'ancrage de l'ARNm sur la 40S *via* son interaction avec le complexe eIF4F-PABP. L'absence de coiffe sur les ARNms archéens et par

conséquent, l'absence d'homologues des différents eIF4s, rendrait alors inutile la présence d'un homologue d'eIF3.

Enfin, les sous-unités catalytiques du facteur d'échange eIF2B étant absentes chez les Archées, le mécanisme par lequel aIF2-GDP est recyclé en aIF2-GTP en fin de démarrage reste inconnu. Notamment, cet échange pourrait survenir de manière spontanée.

### 3) Conclusion.

Finalement, de nombreuses questions restent encore posées concernant le mécanisme du démarrage de la traduction chez les Archées et la place qu'il occupe au sein de l'évolution. En effet, la comparaison de ce processus dans les trois domaines du monde vivant est avant tout un formidable outil pour tenter d'imaginer quel était le mécanisme adopté par notre ancêtre commun LUCA. Les différentes données collectées jusqu'à présent ont d'ailleurs permis d'imaginer un modèle de démarrage partant d'un ARNm monocistronique et sans région 5'-UTR dans lequel seuls les deux facteurs universellement conservés IF1-e/aIF1A et IF2-e/aIF5B interviendraient (Londei, 2005). La **Figure 1.27** illustre ce modèle sous deux angles distincts, selon que l'ancêtre d'IF2 serait ou non capable de lier l'ARNt initiateur.



**Figure 1.27 : Modèle de démarrage de la traduction chez l'ancêtre commun LUCA (Londei, 2005).** Dans ce modèle simplifié, seuls les 2 facteurs IF1 et IF2 sont requis. L'interaction ribosome/ARNm pourrait se faire à l'aide du complexe 70S entier, après positionnement de l'ARNt initiateur au site P du ribosome – positionnement qui surviendrait au site P par défaut, le site A étant occupé par le complexe IF1-IF2. La formation de complexes de démarrage avec les petites sous-unités ribosomales dissociées serait alors survenu plus tard au cours de l'évolution, concomitamment à l'apparition de l'interaction SD-ASD. Dans le modèle **A**), l'ancêtre d'IF2 ne possède pas de propriétés de liaison de l'ARNt. Dans ce cas, il est possible que n'importe quel ARNt ait pu jouer le rôle d'ARNt initiateur en fonction du codon de l'ARNm rencontré au site P. La sélection stricte sur un codon AUG serait alors apparue plus tard. Dans le modèle **B**), l'IF2 ancestral est capable de lier l'ARNt initiateur. Dans ce contexte, l'ARNt initiateur ancestral aurait pu être formylé et la formylation aurait été perdue par la suite chez les Eucaryotes et chez les Archées.

Tout aussi séduisantes soient-elles, ces hypothèses nécessitent encore un immense travail expérimental pour tenter d'apporter des éléments de réponse quant aux mécanismes précis survenant aujourd'hui dans le monde vivant et à plus lointaine échelle de temps, ayant eu lieu chez le progénote. En particulier, de nombreuses étapes clés du processus de démarrage sont encore très mal comprises chez les Eucaryotes.

Dans ce contexte, un facteur de démarrage revêt un intérêt particulier : le facteur hétérotrimérique e/aIF2. Identifié il y a plus d'une trentaine d'années (Picciano *et al.*, 1973 ; Barrieux and Rosenfeld, 1977), e/aIF2 est composé de trois sous-unités, notées  $\alpha$ ,  $\beta$  et  $\gamma$ , qui sont relativement bien conservées au sein des domaines eucaryotique et archéen. La délétion de l'un ou l'autre des gènes codant pour chacune des sous-unités du facteur est létale chez *Saccharomyces cerevisiae* (Donahue *et al.*, 1988 ; Cigan *et al.*, 1989 ; Hannig *et al.*, 1993). Ce complexe de 124 kDa chez la levure et de 93 kDa chez l'archée *P. abyssi* ne possède pas d'orthologue bactérien malgré le rôle central qu'il joue dans l'étape de démarrage. Il est en effet responsable du recrutement de l'ARNt initiateur et de son acheminement sur la petite sous-unité ribosomale mais selon un tout autre mécanisme que le facteur IF2, qui assure une fonction similaire chez les Bactéries. Le travail de thèse présenté dans ce manuscrit s'est donc centré sur ce facteur atypique pour essayer de comprendre son mécanisme d'action spécifique et quelles ont pu être les raisons de sa mise en place au sein d'un système plus évolué tel que le système eucaryotique. Dans le **Chapitre II** de ce manuscrit, nous nous attacherons à faire le point sur les connaissances bibliographiques dont nous disposons au début de ce travail de thèse pour chacune des trois sous-unités du facteur e/aIF2, d'un point de vue structural et fonctionnel.

*Chapitre II : le Facteur  
d'Initiation e/aIF2  
Fonctionnalités et Organisation  
Structurale*



## **Partie A : Rôle général du facteur de démarrage e/aIF2.**

### **1) Description fonctionnelle.**

Chez les Eucaryotes et chez les Archées, e/aIF2 résulte de l'association stœchiométrique de trois sous-unités :  $\alpha$ ,  $\beta$  et  $\gamma$ . Au cours du démarrage de la traduction, ce facteur transporte le Met-ARNt<sub>i</sub><sup>Met</sup> sous forme d'un complexe ternaire avec une molécule de GTP sur la petite sous-unité ribosomale.

e/aIF2 lié au GTP et au Met-ARNt<sub>i</sub><sup>Met</sup> est impliqué dans la reconnaissance du bon codon initiateur. L'hydrolyse du GTP survient lorsque l'appariement correct codon-anticodon est réalisé. Des mutations au sein de chacune de ses trois sous-unités affectent la sélection de l'AUG et permettent le démarrage sur un codon UUG (Donahue *et al.*, 1988 ; Cigan *et al.*, 1989 ; Williams *et al.*, 1989 ; Castilho-Valavicius *et al.*, 1990 ; Dorris *et al.*, 1995).

Au sein de l'hétérotrimère, la sous-unité  $\gamma$  comporte un domaine G identifié par homologie de séquences avec d'autres protéines G impliquées dans la traduction, notamment le facteur d'élongation EF1A (Erickson and Hannig, 1996). Des mutations au niveau de la séquence consensus NKXD rencontrée au sein d'e/aIF2 $\gamma$  affectent très fortement la liaison du GTP/GDP. La sous-unité  $\gamma$  pourrait donc jouer un rôle prépondérant dans la liaison des nucléotides G bien que l'intervention de l'une ou l'autre des sous-unités  $\alpha$  et  $\beta$  ne puisse être exclue (Naranda *et al.*, 1995 ; Erickson and Hannig, 1996).

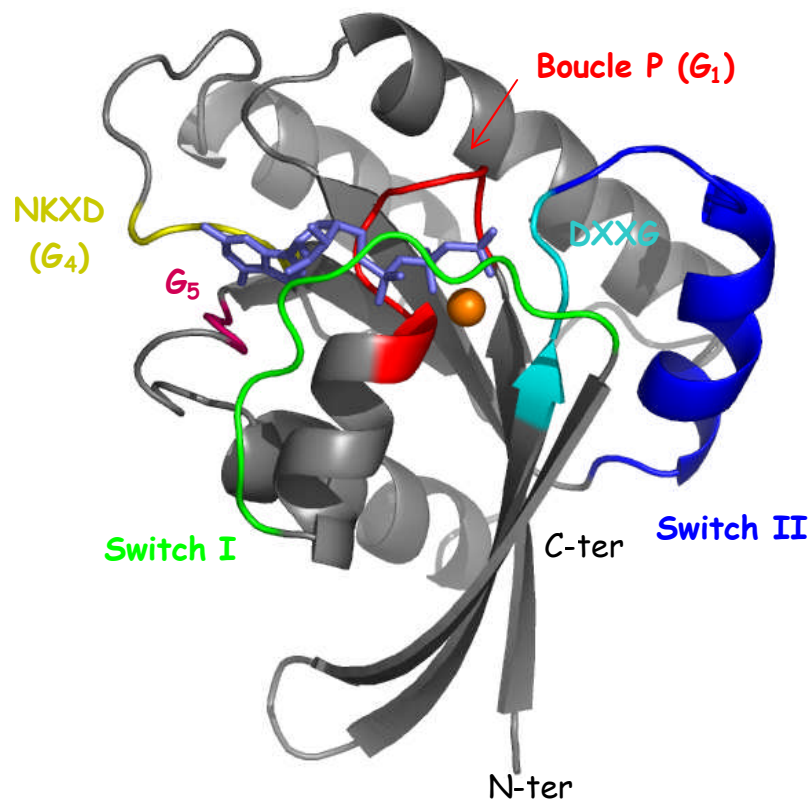
Avant d'entrer plus en détails dans l'étude structurale de la sous-unité  $\gamma$ , nous ferons ici un bref rappel du mécanisme d'action des GTPases (Vetter and Wittinghofer, 2001 ; Sprang, 1997 ; pour revues).

### **2) Les protéines G : un mécanisme universel.**

Au sein des cellules, les protéines G oscillent entre un état conformationnel actif (état ON) et un état conformationnel inactif (état OFF), ce qui leur permet de contrôler une grande variété de processus allant de la croissance et de la différenciation cellulaire au transport nucléo-cytoplasmique et vésiculaire. L'activation des protéines G (passage au mode ON) nécessite la dissociation préalable du nucléotide GDP qui leur est complexé. Ce processus relativement lent doit être accéléré par un facteur noté GEF (pour Guanine Exchange Factor) qui permet l'échange GDP  $\rightarrow$  GTP sur la protéine G. Il s'agit d'une étape généralement réversible. A l'inverse, le passage au mode OFF est un processus intrinsèquement irréversible qui correspond à l'hydrolyse du GTP en GDP – soit la réaction GTPasique à proprement

parler – et ce processus également relativement lent doit être accéléré par une GAP (GTPase Activating Protein).

Les cellules eucaryotiques contiennent entre 100 et 150 protéines G différentes. Celles-ci sont regroupées en plusieurs superfamilles : les protéines dérivées de la protéine Ras, les protéines G hétérotrimériques, les facteurs impliqués dans la synthèse protéique et diverses autres sous-familles. Malgré des différences notables au sein de ces familles, le domaine G de ces protéines présente une structure (**Figure 2.1**) et un mécanisme de changement conformationnel universels. Ainsi, un domaine minimal d'environ 20 kDa, commun à toutes les protéines G, assure les fonctions de liaison et d'hydrolyse des nucléotides guanine. Cette structure minimale comprend 6 brins  $\beta$  et 5 hélices  $\alpha$ . Le domaine G présente cinq séquences consensus impliquées dans sa fonction, à savoir les motifs  $G_1$  :  $GX_4GK(S/T)$  ou boucle P,  $G_2$  : boucle effectrice ou Switch I,  $G_3$  :  $DXXG$  ou Switch II,  $G_4$  :  $(N/T)KXD$  et  $G_5$  :  $(T/G)(C/S)A$  où X représente un résidu quelconque (Sprang, 1997) (Bourne *et al.*, 1991). Autour de ce cœur fonctionnel et universel, les divers domaines G présentent des insertions ou des modifications caractéristiques de chaque protéine et importantes pour leur fonction biologique propre.

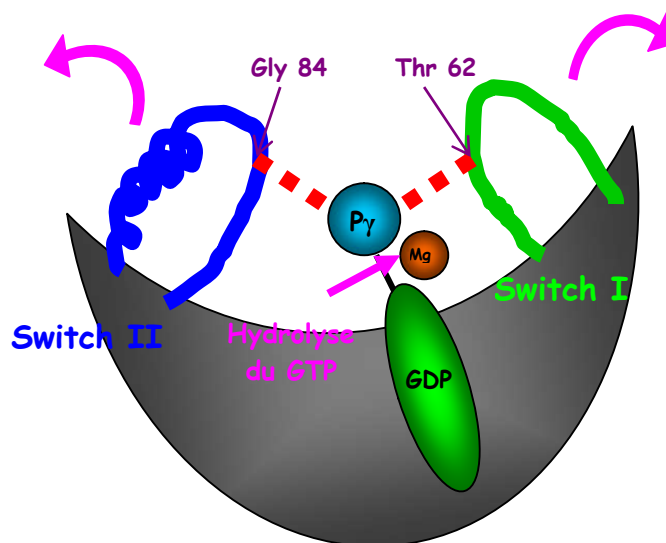


**Figure 2.1 :** Structure du domaine G minimal commun à toutes les protéines G d'après Vetter and Wittinghofer, 2001. Les différentes séquences consensus du domaine G sont représentées sur ce domaine modèle dérivé de la structure du domaine G de la protéine Ras (Pai *et al.*, 1990). L'ion magnésium est représenté par une sphère orange et le nucléotide GTP est représenté en violet.

### 1. Mécanisme du changement conformationnel.

Le changement conformationnel qui intervient au sein du domaine G lors du passage GDP → GTP ou GTP → GDP fait intervenir principalement deux zones de ce cœur structural : les motifs Switch I et Switch II. Les réarrangements qui s'ensuivent alors au sein de la protéine G varient d'une protéine à l'autre et peuvent être relativement importants. Cependant, le mécanisme de base est similaire chez toutes ces GTPases.

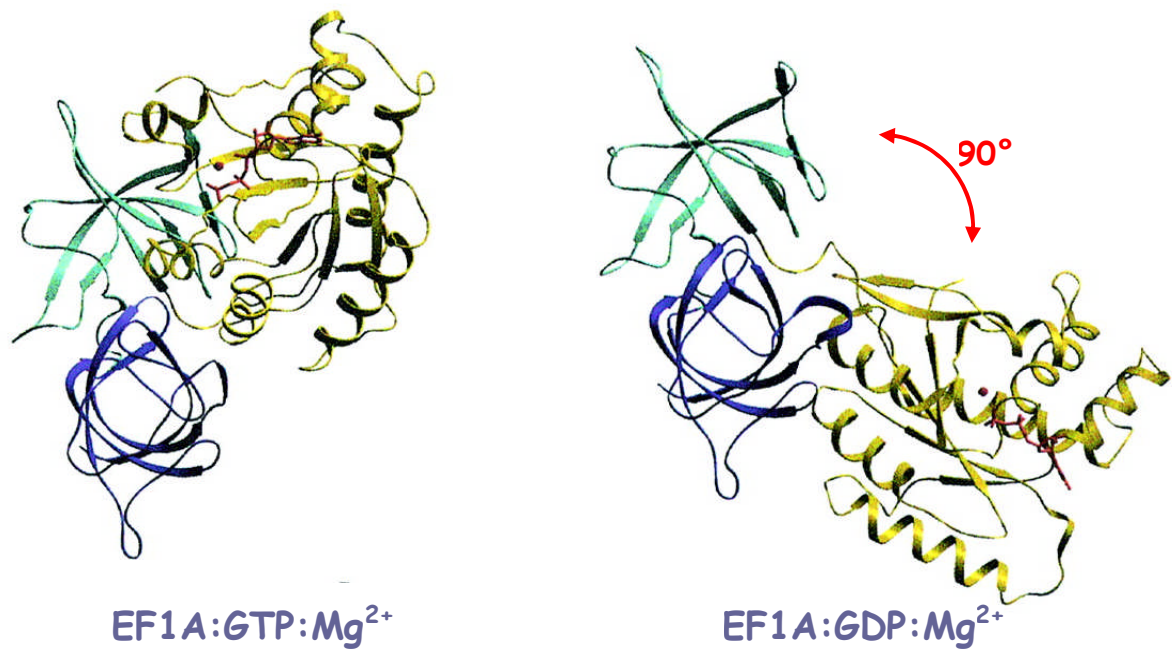
Dans la forme liée au GTP, deux résidus invariants forment une liaison H avec le phosphate  $\gamma$  du GTP par l'intermédiaire du groupe NH de leur chaîne principale. Il s'agit d'une thréonine du motif Switch I (Thr62 chez EF1A) et de la glycine du motif DXXG (Switch II, Gly84 chez EF1A). La thréonine est également impliquée dans la coordination de l'ion  $Mg^{2+}$ . Le départ du phosphate  $\gamma$  après hydrolyse du GTP entraîne la relaxation de ces deux zones Switch du fait de la rupture de ces liaisons H. La **Figure 2.2** schématise ce mouvement de relaxation.



**Figure 2.2 :** Représentation schématique du mouvement conformationnel universel qui survient au sein du domaine G des GTPases lors de l'hydrolyse du GTP en GDP d'après Vetter and Wittinghofer, 2001. La numérotation des invariants Thr et Gly correspond à la protéine EF1A.

Ce mouvement minimal est susceptible d'exister dans toutes les protéines G. Il peut ensuite se répercuter au sein du polypeptide à plus ou moins longue distance. Ainsi, chez EF1A, cette transition impliquant les résidus Thr62 et Gly84 s'accompagne de changements conformationnels majeurs avec un repositionnement du domaine 1 par rapport aux domaines

2 et 3. Le domaine G subit en effet une rotation de 90° par rapport aux tonneaux  $\beta$  et dans cette conformation «ouverte», le site de liaison de l'ARNt n'existe plus (Nissen *et al.*, 1995 ; Berchtold *et al.*, 1993). La **Figure 2.3** représente les deux conformations ON et OFF d'EF1A.



**Figure 2.3 :** Réorganisation du facteur d'élongation EF1A après hydrolyse du GTP en GDP d'après Berchtold *et al.*, 1993.

## **2. Hydrolyse du GTP, GAP-assistée.**

Pour la plupart des protéines G, la réaction d'hydrolyse du GTP est particulièrement lente et ne pourrait pas assurer une fonction biologique convenable car cette dernière nécessite généralement un passage rapide de la protéine à un état inactivé. Ainsi, les protéines GAP sont-elles présentes pour activer cette réaction d'hydrolyse. Le mécanisme exact par lequel la GAP intervient reste encore incertain mais il a pu être précisé grâce à l'obtention de structures de complexes protéine G – GAP, en particulier le complexe Ras - RasGAP (Scheffzek *et al.*, 1997).

De manière générale, la protéine G présente un résidu catalytique invariant au sein du Switch II qui est une glutamine dans le cas de Ras (Gln61). Chez les facteurs de traduction, cette glutamine est remplacée par une histidine (His85 chez EF1A). Cette glutamine (histidine) permet le positionnement correct de la molécule d'eau qui va réaliser l'attaque nucléophile sur le groupe phosphate  $\gamma$  du GTP.

Cependant, cette glutamine n'est pas suffisante pour déclencher l'hydrolyse du GTP et un autre résidu catalytique est requis. Ce second résidu est apporté par la GAP. Elle insère en effet un résidu arginine sous forme d'un motif « Arginine Finger » au niveau du site catalytique. Cette boucle catalytique est elle-même stabilisée par un second résidu invariant de la GAP qui peut être une arginine ou une lysine.

Il faut noter que dans le cas d'EF1A, il n'existe pas de GAP. C'est le motif L7/L12 du ribosome qui active la fonction GTPasique d'EF1A.

### **3. Recyclage GDP → GTP, GEF-assisté.**

Le relargage du GDP des protéines G est un processus relativement lent et la présence d'une GEF permet de l'accélérer de plusieurs ordres de grandeur (environ 5). Le mécanisme d'action de la GEF passe par une série de réactions rapides et réversibles qui correspondent au passage successif du complexe binaire GTPase-GDP à un complexe ternaire GTPase-GDP-GEF puis à un nouveau complexe binaire GTPase-GEF où la protéine G n'est plus liée à aucun nucléotide guanine. Cette série de réactions est alors réinversée par liaison d'un nouveau nucléotide qui est généralement le GTP, du fait de sa présence cellulaire prédominante par rapport au GDP (Cherfils and Chardin, 1999).

La structure de complexes binaires entre la GTPase libérée de tout nucléotide et sa GEF a été obtenue dans le cas de différentes protéines G et en particulier pour le facteur d'élongation 1A (complexe EF-Tu:EF-Ts : Wang *et al.*, 1997 ; complexe eEF1A-eEF1B : Andersen *et al.*, 2000). Bien que les mécanismes diffèrent d'une GTPase à l'autre, quelques caractéristiques communes ont pu être tracées. Ainsi, les GEFs interagissent avec les régions Switch I et II de leur GTPase et en particulier, elles intercalent des résidus à proximité de la boucle GKT (motif G<sub>1</sub>) et de l'ion magnésium. Ces insertions provoquent un changement structural local au sein de la GTPase qui a un effet inhibiteur sur la liaison des phosphates du nucléotide guanine et de l'ion Mg<sup>2+</sup> et permet ainsi le relargage du GDP.

## **Partie B : La sous-unité $\gamma$ du facteur e/aIF2.**

e/aIF2 $\gamma$  est la sous-unité principale du facteur d'initiation 2. Elle est codée par le gène essentiel *GCD11* (Hannig *et al.*, 1993). La protéine de levure contient 527 résidus (soit 55 kDa). Les protéines archéennes ont généralement un poids moléculaire plus faible (autour de 45 kDa). Cette différence s'explique par la présence d'une extrémité N-terminale additionnelle chez les Eucaryotes, de longueur variable et pouvant atteindre 90 résidus, et pour laquelle aucun rôle précis n'a été identifié jusqu'à présent (voir alignement de séquences en **Annexe 1**). La sous-unité  $\gamma$  constitue le cœur de l'hétérotrimère e/aIF2. Ainsi, les deux autres sous-unités  $\alpha$  et  $\beta$  interagissent avec la protéine  $\gamma$ . Par contre, elles ne s'associent pas entre elles (Schmitt *et al.*, 2002).

### **1) Structure de la sous-unité $\gamma$ .**

La structure cristallographique de la protéine e/aIF2 $\gamma$  a été résolue pour la première fois dans le cas de l'archée hyperthermophile *Pyrococcus abyssi* à 1.8 Å de résolution (Schmitt *et al.*, 2002) (**Figure 2.5**).

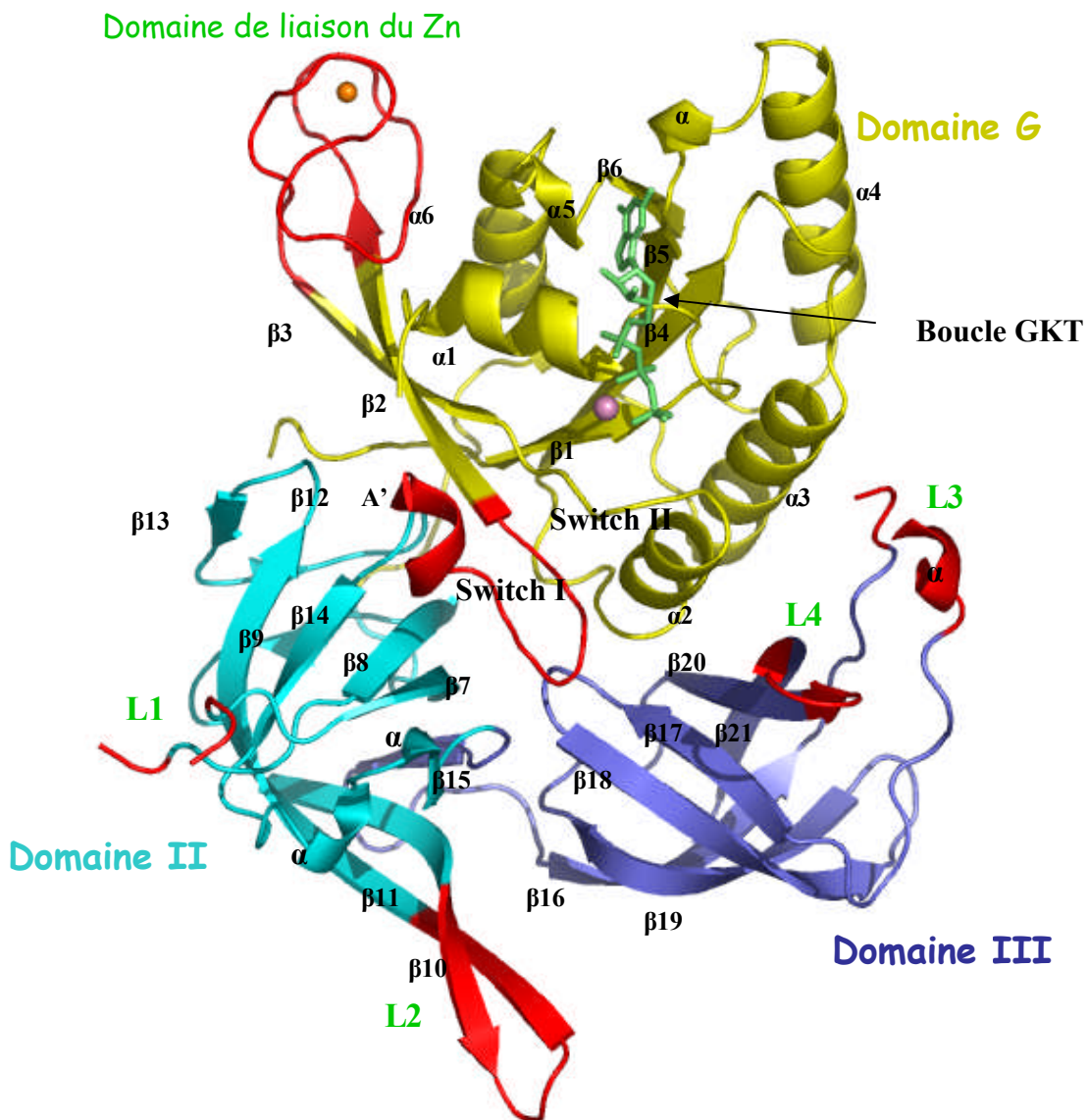
La protéine est organisée en trois domaines. Le domaine G (résidus 1 à 205) correspond au domaine de fixation des nucléotides guanine. Les domaines II (résidus 206 à 308) et III (résidus 309 à 410) présentent tous deux un repliement en tonneau  $\beta$ .

Les séquences consensus des protéines G sont retrouvées dans la structure d'aIF2 $\gamma$ . Le mode de liaison des nucléotides G à e/aIF2 $\gamma$  a pu être étudié après trempage de cristaux de protéine libre dans des solutions contenant les nucléotides GDP-Mg<sup>2+</sup> ou GDPNP-Mg<sup>2+</sup> à 10 mM (Schmitt *et al.*, 2002). Ainsi, la poche de liaison des nucléotides guanine d'aIF2 $\gamma$  contient les motifs suivants : la boucle **GKT** (résidus 22 à 24), la séquence **QNKIE** (résidus 144 à 148) reliant le brin  $\beta$ 5 à, l'hélice  $\alpha$  et la séquence **SALH** (résidus (180 à 183) qui correspond à la petite hélice  $\alpha$ 5. Ces deux dernières régions assurent la spécificité vis-à-vis de la base guanine. L'ion magnésium est hexacoordiné par la thréonine 24 de la boucle GKT, par les phosphates  $\beta$  et  $\gamma$  du GTP et par 3 molécules d'eau. Sous forme GDP, la coordination du phosphate  $\gamma$  du GTP est remplacée par celle d'une quatrième molécule qui interagit également avec la chaîne latérale de l'aspartate 89.

De manière générale, Pa-aIF2 $\gamma$  présente une organisation structurale similaire à celle des facteurs d'élongation et en particulier le facteur bactérien EF1A (Berchtold *et al.*, 1993 ; Nissen *et al.*, 1995). Ainsi, la superposition de chacun des trois domaines de  $\gamma$  avec les domaines correspondants au sein d'EF1A:GDPNP conduit aux rmsd suivants :

Domaines comparés	Nombres de Carbones C $\alpha$ comparés	Rmsd en Å
Domaine G	138	1,74
Domaine II	89	1,65
Domaine III	74	1,50

**Tableau 2.4 :** Rmsd mesurés entre les domaines d'aIF2 $\gamma$  de *P. abyssi* et les domaines correspondants de son homologue structural EF1A en conformation GDPNP d'après Schmitt *et al.*, 2002.



**Figure 2.5 :** Structure de la protéine aIF2 $\gamma$  de l'archée *P. abyssi* complexée à du GDPNP à 1.8 Å de résolution (Schmitt *et al.*, 2002). Le nucléotide GDPNP est représenté en vert complexé à un ion Mg<sup>2+</sup> en rose. Un ion Zn<sup>2+</sup> (en orange) est également ligandé par 4 cystéines dans le domaine additionnel (comparativement à EF-Tu) qui pointe hors du domaine G. Les éléments de structure en rouge indiquent les zones de divergence par rapport à EF1A.

Dans le cas d'EF1A, un large mouvement du domaine G comparativement aux domaines II et III survient entre les formes GDP et GTP du facteur d'élongation (La Cour *et al.*, 1985 ; Berchtold *et al.*, 1993 ; Polekhina *et al.*, 1996) (**Figure 2.5**). Au contraire, l'orientation relative des trois domaines d'aIF2 $\gamma$  est identique au sein des trois structures de la sous-unité (protéine sans nucléotide G, protéine liée au GDP et protéine liée au GTP). Dans ces trois états nucléotidiques,  $\gamma$  adopte une conformation fermée pour laquelle le domaine G reste à proximité des domaines II et III. La superposition des domaines II et III d'aIF2 $\gamma$  par rapport aux domaines correspondants du facteur d'élongation au sein de la structure EF1A:GDPNP (Nissen *et al.*, 1995) révèle une rotation de 14° entre les domaines G des deux protéines (Schmitt *et al.*, 2002). Cependant, un déplacement de 84° est obtenu lorsque la superposition est réalisée avec la forme EF1A:GDP du facteur d'élongation. Ainsi, quel que soit son état nucléotidique, aIF2 $\gamma$  adopte une conformation plus proche de celle observée pour la forme « active » EF1A:GDPNP du facteur d'élongation que celle de la forme « inactive » EF1A:GDP.

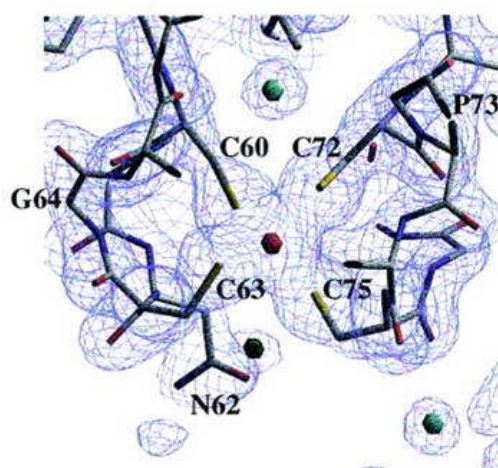
## **2) Zones spécifiques au facteur de démarrage.**

Bien que Pa-aIF2 $\gamma$  et EF1A soient très proches d'un point de vue structural, les alignements de séquence et la superposition des structures de Pa-aIF2 $\gamma$  (Schmitt *et al.*, 2002) et d'EF1A:GDPNP de *Thermus aquaticus* (Nissen *et al.*, 1995) ont permis de mettre en évidence des zones qui diffèrent entre les deux facteurs (zones indiquées en rouge sur la **Figure 2.5**).

La première différence notable concerne le domaine 2 d'aIF2 $\gamma$ , qui possède deux boucles qui ne sont pas présentes chez EF1A. Il s'agit tout d'abord de la boucle L1 entre les brins  $\beta 7$  et  $\beta 8$ , qui englobe les résidus 220 à 228 (numérotation *P. abyssi*). Cette boucle n'est pas totalement visible sur la structure de la **Figure 2.5** car la région correspondant aux résidus 224 à 227 était désordonnée et par conséquent non interprétable sans ambiguïté dans la densité électronique. La seconde zone caractéristique est la boucle en épingle à cheveux L2 entre les brins  $\beta 10$  et  $\beta 11$ , qui contient les résidus 255 à 265 (numérotation *P. abyssi*). Ces deux boucles se situent dans le domaine 2 de Pa-aIF2 $\gamma$  qui présente un motif de repliement de type tonneau  $\beta$  et elles dépassent très nettement de la structure.

Une autre région caractéristique de Pa-aIF2 $\gamma$  par rapport à EF1A correspond aux résidus 340-343. Cette insertion est toujours rencontrée dans les séquences d'e/aIF2 $\gamma$ . Une autre boucle reliant les brins  $\beta 18$  et  $\beta 19$  est raccourcie chez e/aIF2 $\gamma$  par rapport au facteur d'élongation. Enfin, une légère extension de la boucle reliant les brins  $\beta 20$  et  $\beta 21$  dans le domaine 3 est également observée chez Pa-aIF2 $\gamma$ .

Pour finir, la structure de Pa-aIF2 $\gamma$  met en évidence l'existence d'un domaine de fixation du zinc inséré dans le domaine G (résidus 57 à 81 selon la numérotation de *P. abyssii*). Chez EF1A, la position équivalente à ce domaine est occupée par la partie C-terminale de l'hélice  $\alpha 1$ . Au sein de ce domaine, l'ion Zn<sup>2+</sup> est chélaté par quatre cystéines (C60, C63, C72 et C75) (**Figure 2.6**). La mutation de l'une des ces cystéines chez la levure entraîne un sévère défaut de croissance (Erickson *et al.*, 1997). Ceci suggère un rôle fonctionnel important de ce domaine.

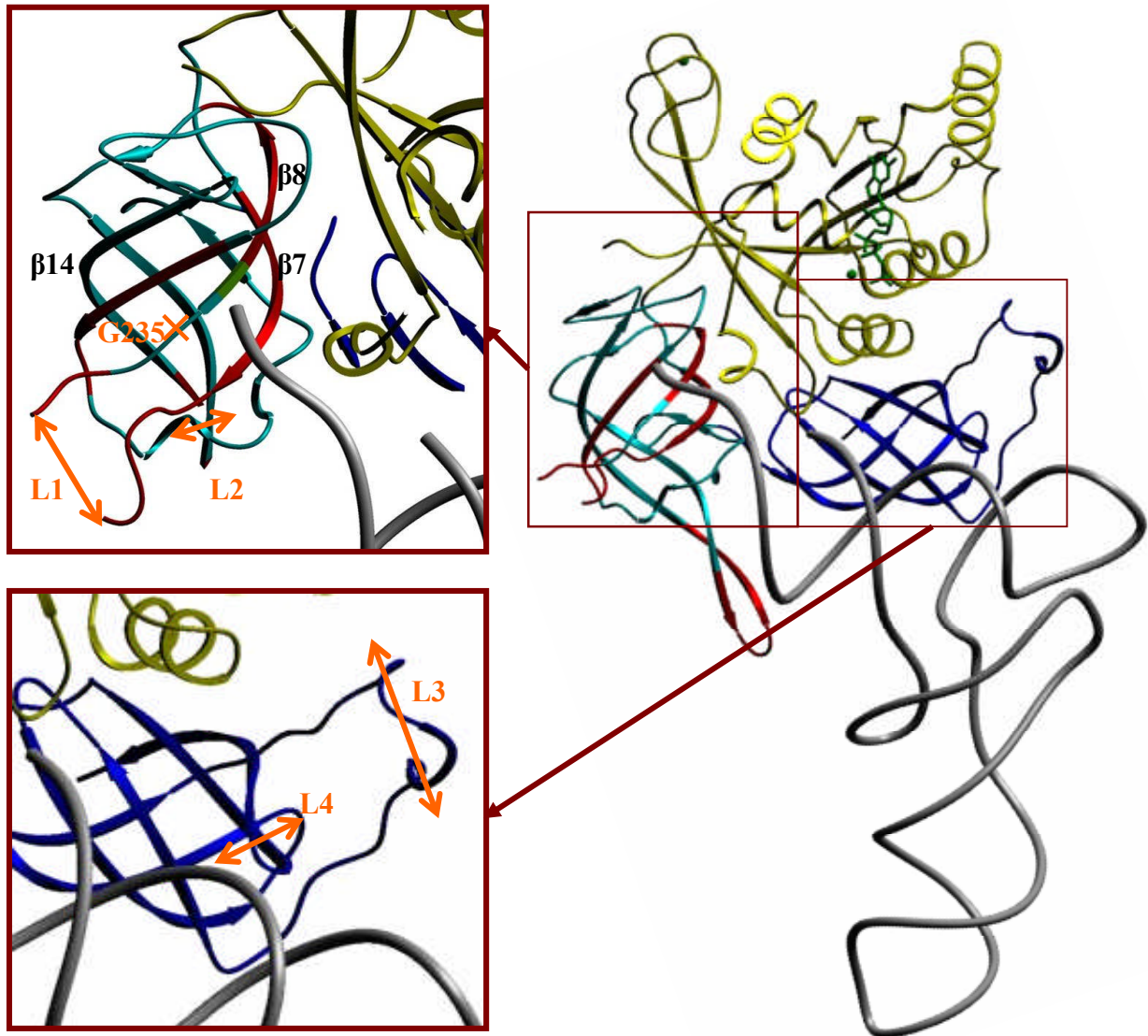


**Figure 2.6 :** Tétracoordination de l'ion Zn<sup>2+</sup> par 4 cystéines au sein de la sous-unité aIF2 $\gamma$  de *P. abyssii* (Schmitt *et al.*, 2002).

### 3) Modèle de positionnement de l'ARNt.

La ressemblance structurale entre aIF2 $\gamma$  et EF1A a permis d'envisager un mode de reconnaissance de la molécule d'ARNt similaire au sein des deux protéines. Cette hypothèse a conduit à un modèle de positionnement de l'ARNt sur la sous-unité  $\gamma$ , déduit de la structure du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995) qui est représenté sur la **Figure 2.7** (Schmitt *et al.*, 2002). Le positionnement de l'ARNt sur aIF2 $\gamma$  a été réalisé en superposant les domaines 2 et 3 des deux protéines aIF2 $\gamma$  et EF1A (Schmitt *et al.*, 2002).

Dans le modèle ainsi obtenu, peu de contacts défavorables sont observés. Le plus notable concerne la chaîne latérale de la phénylalanine portée par le Phe-ARNt<sup>Phe</sup> et la chaîne latérale du résidu L290. Cependant, il n'est pas en désaccord avec le modèle de positionnement puisque dans le cas du facteur de démarrage, c'est un résidu méthionine moins encombrant qui serait inséré dans cette poche de fixation.



**Figure 2.7 :** Modèle de positionnement du Phe-ARNt<sup>Phe</sup> sur la structure d'aIF2γ-GDPNP de *P. abyssi* à partir de la structure du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995). Les deux zones potentielles d'ancrage de l'ARNt sur γ ont été agrandies sur la partie gauche de la figure. Dans le domaine 2 d'aIF2γ, une poche formée par les brins β7, β8 et β14 pourrait fixer l'extrémité aminoacylée de la tige acceptrice de l'ARNt, tandis que le bras TΨC interagirait avec le domaine 3 de γ.

Dans le complexe EF1A:GTP:ARNt, deux zones d'interaction sont identifiables (Nissen *et al.*, 1995). La tige TΨC de l'ARNt est en contact avec le domaine 3 d'EF1A. Cependant aucune interaction spécifique avec le polynucléotide n'est détectée, ce qui est en accord avec la grande variété d'ARNts, et donc de séquences nucléotidiques, que peut fixer EF1A. Le second point de contact correspond à la zone d'interaction avec la tige acceptrice de l'ARNt. Celle-ci s'engouffre dans un tunnel délimité d'un côté par la région effectrice et de l'autre côté par le domaine 2 d'EF1A. La base 76 et l'acide aminé s'insèrent quant à eux dans une large poche délimitée par les résidus D227, E226, F229 et T239 (résidus R215, L214,

F217 et G235 chez aIF2 $\gamma$  de *P. abyssii*). Cette poche est suffisamment large pour pouvoir s'adapter aux différents acides aminés portés par les ARNts élongateurs.

Ce modèle de positionnement implique qu'une région délimitée par les brins  $\beta 7$ ,  $\beta 8$  et  $\beta 14$  du domaine 2 d'e/aIF2 $\gamma$  constituerait la poche de fixation du résidu méthionine de l'ARNt initiateur. D'autre part, les résidus Phe52 et Gly45 (qui correspondent aux résidus Tyr142 et Asn135 chez la levure) sont en bonne position au niveau des régions Switch I et II pour interagir avec le polynucléotide. Chez la levure, les mutations Y142H et N135K affectent la capacité de liaison à l'ARNt de l'hétérotrimère eIF2 (Huang *et al.*, 1997). Enfin, la face latérale du tonneau  $\beta$  du domaine III d'aIF2 $\gamma$  pourrait interagir avec le bras T $\psi$ C de l'ARNt. La validation de ce modèle n'a toutefois pas encore été effectuée.

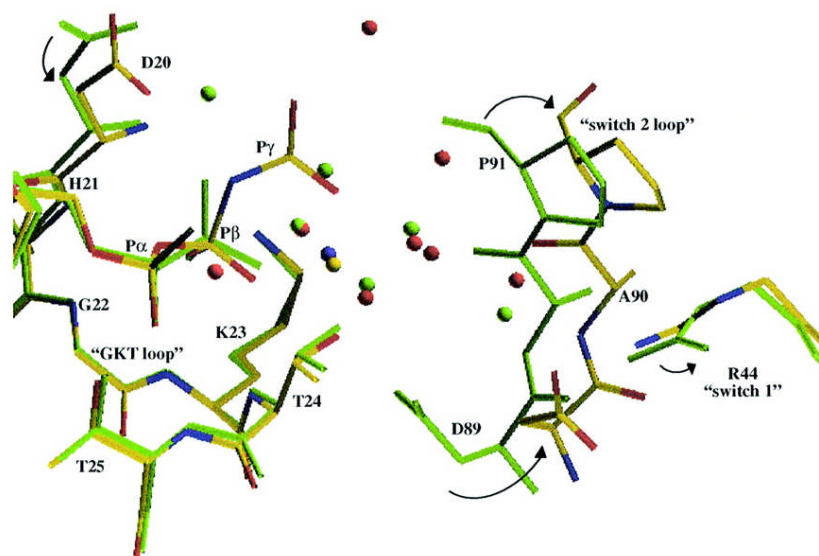
Chez EF1A, le groupe 5'-phosphate de l'ARNt est impliqué dans des interactions fortes avec trois résidus universellement conservés chez les facteurs d'élongation : K90, N91 et R300, selon la numérotation d'EF1A. Chez e/aIF2 $\gamma$ , ces résidus sont remplacés par des résidus non chargés et moins bien conservés (résidus T98, T99 et A308 chez aIF2 $\gamma$  de *P. abyssii*). Ceci suggère que le groupe 5'-phosphate du Met-ARNt<sub>i</sub><sup>Met</sup> n'est pas reconnu par aIF2 $\gamma$ . De plus, la région 271-290 qui, chez EF1A assure des contacts avec le groupe  $\alpha$ -amine et la liaison ester du Phe-ARNt<sup>Phe</sup>, présente une conformation différente dans la structure de Pa-aIF2 $\gamma$ . Ces caractéristiques propres au facteur de démarrage pourraient ainsi contribuer à l'exclusion des ARNts élongateurs dans l'étape de démarrage.

#### **4) Mouvement des régions Switch : comparaison d'e/aIF2 $\gamma$ avec EF1A.**

Dans le cas du facteur d'élongation EF1A, un déplacement caractéristique des régions Switch est observé entre les formes GTP et GDP (Berchtold *et al.*, 1993 ; Polekhina *et al.*, 1996) (**Figure 2.3**). Ce mouvement pourrait être à l'origine du large réarrangement des domaines 2 et 3 par rapport au domaine 1 observé entre les formes GTP et GDP (**Figure 2.3**).

Pour aIF2 $\gamma$ , la liaison du GDP-Mg<sup>2+</sup> entraîne, par rapport à la forme apo de la protéine, une restructuration de la boucle GKT qui est transmise aux résidus voisins interagissant avec ce motif, notamment les résidus de la région Switch II. En particulier, les résidus D89 à P91 se retrouvent dans une conformation qui leur permet d'interagir avec l'ion Mg<sup>2+</sup> par l'intermédiaire de liaisons hydrogène avec les molécules d'eau entourant le magnésium (**Figure 2.8**). Lors de la liaison du GDPNP-Mg<sup>2+</sup>, la boucle GKT est dans la même conformation que dans la protéine liée au GDP. Par contre, la région Switch II retrouve une conformation similaire à celle observée au sein de la protéine sans nucléotide G. En

particulier, la proline P91 subit un déplacement qui permet d'éviter un encombrement stérique avec le phosphate  $\gamma$  du nucléotide G. Le mouvement est répercuté au sein de la région Switch II pour permettre l'accommodation du GDPNP dans la poche de fixation du nucléotide G (**Figure 2.8**). Par contre, aucun mouvement de la région Switch I n'est observé entre les formes GDP et GDPNP d'aIF2 $\gamma$  et dans les deux cas, cette région conserve une conformation similaire à celle observée au sein d'EF1A:GDP. De plus, si l'on retrouve chez aIF2 $\gamma$  de *P. abyssi* les deux résidus Thr47 et Gly92 universellement conservés des GTPases (Thr62 et Gly84 chez EF1A), ceux-ci ne semblent pas impliqués dans l'interaction avec le phosphate  $\gamma$  du GTP et ils ont donc peu de chances d'induire le mouvement des régions Switch classiquement observé au sein des domaines G (Schmitt *et al.*, 2002).



**Figure 2.8 :** Mouvements des régions Switch I et II d'aIF2 $\gamma$  de *P. abyssi* après fixation de GDP-Mg $^{2+}$  ou de GDPNP-Mg $^{2+}$  d'après Schmitt *et al.*, 2002.

Ainsi, comparativement au facteur d'élongation EF1A, aIF2 $\gamma$  ne présente qu'un faible réarrangement des régions Switch lors de la transition GDP  $\rightarrow$  GTP et la structure reste en conformation fermée avec les domaines 2 et 3 en contact avec le domaine 1. Cette conformation fermée est également rencontrée dans la structure d'aIF2 $\gamma$  de l'archée *Methanococcus jannaschii* (Roll-Mecak *et al.*, 2004).

Le mouvement limité des régions Switch observé lors de la transition GDP  $\rightarrow$  GTP ne permet pas d'expliquer simplement la GTP-dépendance de la liaison du Met-ARNt $^{\text{Met}}$  par eIF2 (Hershey and Merrick, 2000). L'absence d'un large mouvement des régions Switch d'aIF2 $\gamma$  observé dans ces structures pourrait cependant s'expliquer par la méthode même

d'obtention des cristaux. En effet, les cristaux de protéine complexée à du GDP ou à du GDPNP ont été obtenus par trempage des cristaux de la protéine sous forme apo dans des solutions contenant les nucléotides G. Des structures identiques sont obtenues à partir de cristaux produits par ensemencement de solutions protéiques contenant le nucléotide avec des microcristaux de la protéine sous forme apo. L'empilement cristallin existant au sein des cristaux de protéine libre pourrait empêcher un réarrangement complet des régions Switch de  $\gamma$ . Toutefois, il est également possible que les changements de conformation d'aIF2 $\gamma$  ne soient complets qu'en présence des deux autres sous-unités  $\alpha$  et  $\beta$  de l'hétérotrimère (Schmitt *et al.*, 2002).

### **5) Perspectives.**

La structure de la sous-unité aIF2 $\gamma$  a révélé une très forte homologie structurale avec le facteur d'élongation EF1A. Cette ressemblance permet d'envisager un mode de reconnaissance de l'ARNt initiateur par e/aIF2 similaire à celui observé pour EF1A. L'un des objectifs majeurs de ce travail de thèse a donc consisté à apporter la validation expérimentale de ce modèle. D'autre part, nous venons de voir qu'aIF2 $\gamma$  possède des caractéristiques propres qui le distinguent du facteur d'élongation. Sont-elles suffisantes pour expliquer le rôle spécifique d'e/aIF2 dans le démarrage de la traduction ? Contrairement à EF1A, e/aIF2 $\gamma$  n'existe pas sous forme monomérique au sein des cellules. La protéine est en effet toujours associée aux deux autres sous-unités  $\alpha$  et  $\beta$ . Dans ces conditions, la question du rôle exact de ces deux autres sous-unités dans la fonction d'e/aIF2 reste posée. En particulier, la GTP-dépendance de la fixation de l'ARNt sur e/aIF2 $\gamma$  s'expliquant difficilement par les faibles changements conformationnels observés au sein de la seule sous-unité  $\gamma$ , e/aIF2 $\alpha$  et e/aIF2 $\beta$  participeraient-elles à la liaison de l'ARNt ?

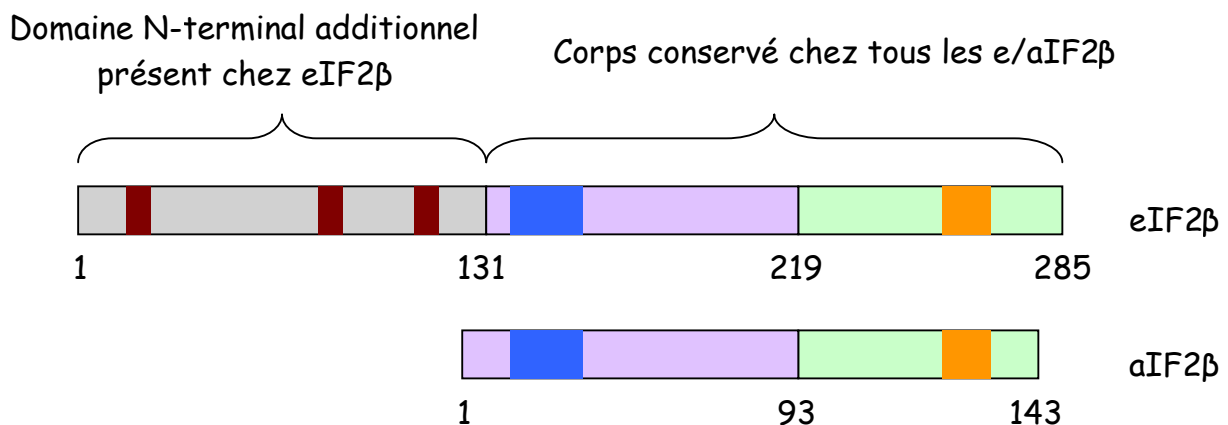
Avant de tenter de répondre à ces questions, nous décrivons les connaissances dont nous disposons à propos des sous-unités  $\alpha$  et  $\beta$  du facteur e/aIF2.

## **Partie C : La sous-unité $\beta$ du facteur e/aIF2.**

### **1) Généralités.**

eIF2 $\beta$  est la seconde sous-unité du facteur d'initiation 2. Elle est codée par le gène essentiel *SUI3*. Cette protéine varie de 250 à 350 résidus selon les espèces (38 kDa chez la levure). Chez les Archées, la protéine est beaucoup plus petite (16 kDa chez *P. abyssi*) car tout le domaine N-terminal est absent (voir alignement de séquences en **Annexe 2**).

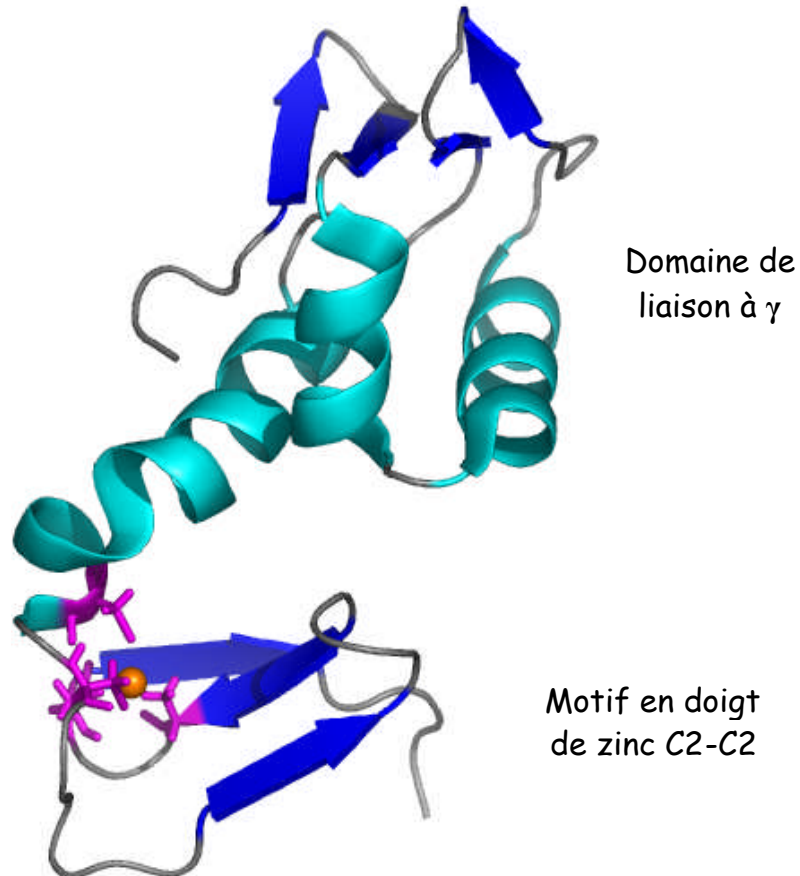
Ainsi, contrairement aux sous-unités  $\alpha$  et  $\gamma$  qui sont très conservées chez les Eucaryotes et chez les Archées, la sous-unité  $\beta$  présente une différence majeure corrélée à la divergence existant entre les mécanismes de démarrage de la traduction dans ces deux domaines du monde vivant. En effet, le domaine N-terminal additionnel présent chez eIF2 $\beta$  et absent chez aIF2 $\beta$  permet la liaison à deux autres facteurs de démarrage : la GAP eIF5 et la GEF eIF2B. Or ces deux facteurs sont également absents chez les Archées, ce qui explique probablement l'absence du domaine N-terminal supplémentaire au sein de la protéine archéenne. Concernant les deux domaines C-terminaux de la protéine communs aux Eucaryotes et aux Archées, une forte homologie est par contre observée entre les différents e/aIF2 $\beta$ . Ces deux domaines sont également très proches du domaine N-terminal de la GAP eucaryotique eIF5 (Thompson *et al.*, 2000). Ceci suggère qu'eIF5 a probablement évolué à partir de ce cœur conservé d'e/aIF2 $\beta$ , apparaissant par duplication de gène pour répondre à de nouveaux besoins dans le processus de démarrage eucaryotique, tandis que la sous-unité eIF2 $\beta$  se dotait d'un domaine N-terminal additionnel (**Figure 2.9**).



**Figure 2.9 :** Organisation structurale d'e/aIF2 $\beta$  (Numérotation *S. cerevisiae* pour eIF2 $\beta$  et de *P. abyssi* pour aIF2 $\beta$ ). Les motifs polylysines du domaine N-terminal additionnel des protéines eucaryotiques sont représentés en marron. La région impliquée dans la liaison à e/aIF2 $\gamma$  est colorée en bleu. Et le motif en doigt de Zinc du domaine C-terminal est représenté en orange.

## 2) Structure de la sous-unité archéenne $\alpha\text{IF}2\beta$ .

La structure de la sous-unité  $\beta$  a été déterminée par RMN pour deux protéines archéennes :  $\alpha\text{IF}2\beta$  de *Methanococcus jannaschii* (Cho and Hoffman, 2002) et  $\alpha\text{IF}2\beta$  de *Methanobacterium thermoautotrophicum* (Gutierrez *et al.*, 2004) (**Figure 2.10**).



**Figure 2.10 :** Structure par RMN de la sous-unité  $\beta$  du facteur  $\alpha\text{IF}2$  de *M. thermoautotrophicum* (Gutierrez *et al.*, 2004). L'ion  $\text{Zn}^{2+}$  tétracoordiné par 4 cystéines universellement conservées est représenté en orange.

La protéine se divise en deux domaines : le domaine N-terminal (résidus 1 à 98) de topologie  $\alpha\beta\beta\alpha\alpha\beta\beta$  qui lie la sous-unité  $\gamma$  et le domaine C-terminal (résidus 99 à 135) qui possède un motif en doigt de zinc.

Les deux domaines d' $\alpha\text{IF}2\beta$  se replient de manière indépendante et ne semblent pas interagir l'un avec l'autre. Ils sont reliés par un polypeptide hélical. Ce dernier a une taille particulièrement bien conservée parmi les différents  $\epsilon/\alpha\text{IF}2\beta$  ce qui suggère qu'il joue un rôle important dans le maintien de l'orientation relative des deux domaines de  $\beta$ .

### **3) Fonction d'e/aIF2 $\beta$ .**

La sous-unité  $\beta$  se lie à  $\gamma$  mais n'interagit pas avec  $\alpha$ . La zone correspondant au domaine central de la protéine eucaryotique - soit les résidus 128 à 205 chez *Saccharomyces cerevisiae* - et par conséquent, au domaine N-terminal de la protéine archéenne, est nécessaire et suffisante à l'interaction avec  $\gamma$  (Thompson *et al.*, 2000). Une étude plus fine a permis de caractériser 9 résidus essentiels pour l'interaction  $\beta$ - $\gamma$  : il s'agit des résidus Tyr131, Ser132, Glu133, Leu134, Leu135, Phe138, Ile141, Leu142 et Leu149 (numérotation de la levure). En effet, la mutation combinée de plusieurs de ces résidus en alanine (double, triple et hexa-mutants Y<sub>131</sub>S<sub>132</sub>, E<sub>133</sub>L<sub>134</sub>L<sub>135</sub> et L<sub>134</sub>L<sub>135</sub>F<sub>138</sub>I<sub>141</sub>L<sub>142</sub>L<sub>149</sub> pour lesquels tous les résidus indiqués sont mutés en alanine) entraîne dans chaque cas l'abolition de la liaison à  $\gamma$  à la fois *in vitro* et *in vivo* (Hashimoto *et al.*, 2002). De plus, la double mutation Y<sub>131</sub>S<sub>132</sub>→AA permet le démarrage de la traduction sur un codon UUG. Ainsi, l'interaction  $\beta$ - $\gamma$  est également un paramètre déterminant dans la sélection du bon codon de démarrage.

Sur le plan fonctionnel, la sous-unité  $\beta$  a d'abord été impliquée dans la liaison du GTP et de l'ARNt initiateur (Gaspar *et al.*, 1994 ; Bommer *et al.*, 1991). Concernant le GTP, c'est finalement la sous-unité  $\gamma$  qui a été reconnue responsable de cette liaison. Cependant, eIF2 $\beta$  possédant deux des séquences consensus de liaison du nucléotide G, il a été proposé que la sous-unité se positionne au voisinage du site de fixation du GTP sur  $\gamma$ . Cette hypothèse est d'ailleurs en accord avec le rôle fonctionnel de la sous-unité puisque chez les Eucaryotes, celle-ci lie eIF5 et eIF2B, tous les deux impliqués dans l'hydrolyse ou le recyclage des nucléotides guanine. Chez les Archées, la partie proche du site de liaison du GTP (c'est-à-dire le domaine N-terminal manquant) serait inexistante. Une autre explication proposée pour expliquer la présence de deux séquences consensus de liaison du GTP est que  $\beta$  lie de l'ATP. Il a été effectivement montré qu'eIF2 pouvait lier d'autres nucléotides tels que l'ATP mais cette liaison inhiberait la formation du complexe ternaire en empêchant l'interaction eIF2-GTP (Flynn *et al.*, 1994 ; Gonsky *et al.*, 1992).

L'implication de la sous-unité  $\beta$  dans la liaison à l'ARNt est pour le moins controversée. En effet, des expériences de pontage chimique ont impliqué cette sous-unité dans une interaction avec l'ARNt. eIF2 $\beta$  se retrouverait ainsi à proximité du polynucléotide au sein du complexe ternaire (Nygard *et al.*, 1980). Cependant, d'autres études ont suggéré qu'en l'absence de la sous-unité  $\beta$ , l'hétérodimère  $\alpha\gamma$  était parfaitement actif dans la formation du complexe ternaire (Colthurst and Proud, 1986 ; Mitsui *et al.*, 1981). Ces résultats ont toutefois été considérés avec prudence car les préparations d' $\alpha\gamma$  ont été obtenues par protéolyse ménagée d'hétérotrimères eIF2 et elles pourraient encore contenir des traces d'eIF2 $\beta$  suffisantes pour assurer la fonction biologique de la sous-unité *in vivo*. De plus,

certaines de ces préparations ne lient plus l'ARNt (Flynn *et al.*, 1993). Ainsi, l'intervention de  $\beta$  dans la liaison du Met-ARNt<sup>Met</sup> reste incertaine.

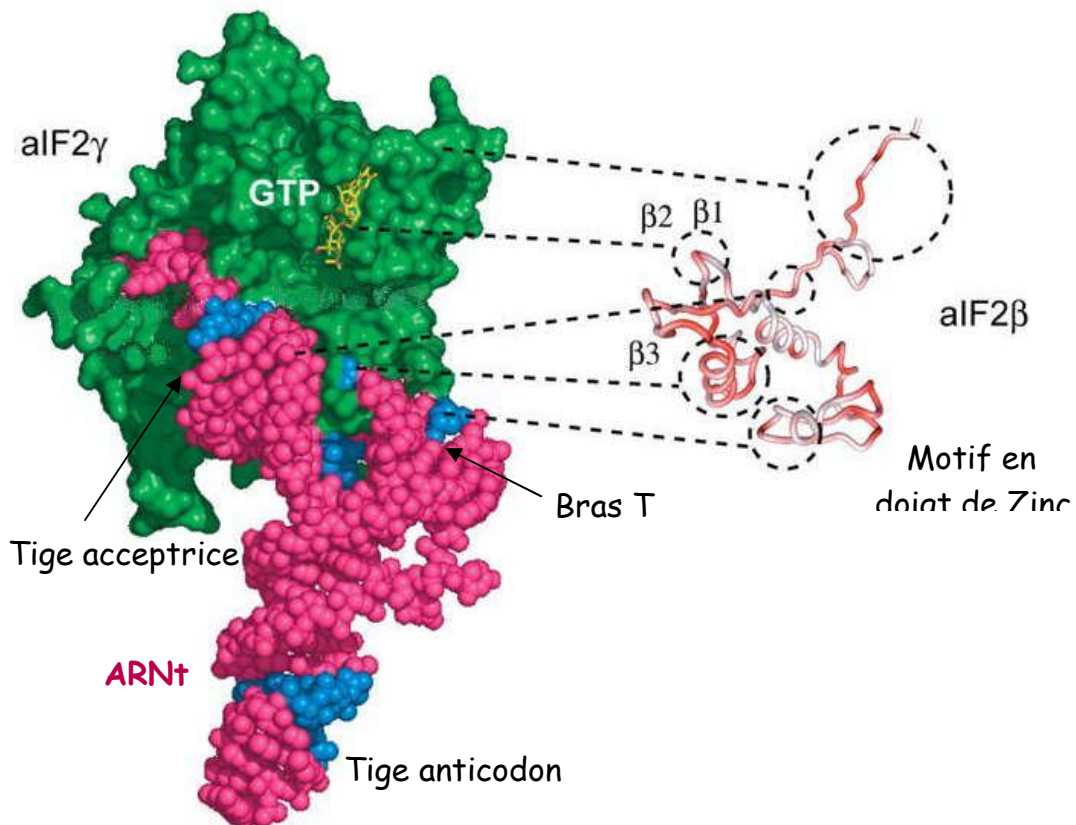
Enfin, la sous-unité  $\beta$  pourrait aussi s'associer à l'ARNm (Flynn *et al.*, 1994 ; Gonsky *et al.*, 1992). Elle possède en effet un site capable de lier des ARNs *in vitro* (Laurino *et al.*, 1999). Ce site, identifié par pontage chimique, correspond au motif C<sub>2</sub>-C<sub>2</sub> en doigt de zinc présent dans le domaine C-terminal de la protéine. La présence d'un atome de Zn est essentielle à l'intégrité structurale et fonctionnelle de la protéine. En l'absence de Zn, l'extrémité C-terminale de la protéine est incorrectement repliée ce qui entraîne son inactivation (Gutierrez *et al.*, 2002). De plus, l'altération d'un seul des quatre résidus cystéines de ce motif ou bien une variation au sein de leur espacement provoque l'inactivation du facteur (Castilho-Valavicius *et al.*, 1992). Le zinc est donc un composant essentiel pour la stabilité de la protéine.

Le motif en doigt de zinc a été soumis à une étude mutationnelle intensive qui a révélé son implication dans la sélection du bon codon de démarrage (Castilho-Valavicius *et al.*, 1990 ; Donahue *et al.*, 1988). Ainsi, une modification des résidus Arg248, Arg253, Leu254, Ser264, Ser267 ou Val268 permet un démarrage de la traduction sur un codon UUG. De manière intéressante, les mutants S264Y et L254P permettent de conférer au facteur eIF2 une activité GTPasique intrinsèque plus élevée et indépendante d'eIF5. Ainsi, le démarrage sur un codon UUG pourrait s'expliquer par un relargage prématuré de l'ARNt initiateur. Dans tous les cas, cette région d'e/aIF2 $\beta$  est cruciale pour la reconnaissance du bon codon initiateur.

Chez les Eucaryotes, un autre site potentiel d'interaction avec l'ARNm a été identifié au sein du domaine N-terminal additionnel, au niveau de motifs polylysines répétitifs. En effet, un mutant d'eIF2 $\beta$  ne possédant plus les trois motifs polylysines, qui contiennent chacun 7 résidus lysines, est incapable de lier le polynucléotide. Par contre, il reste parfaitement actif dans la formation du complexe ternaire et dans son interaction avec la 40S. Ceci suggère que si ces motifs lient un ARN, ce n'est ni l'ARNt ni l'ARNr. Une étude plus poussée a montré que la présence d'un seul bloc polylysines était requise pour assurer une fonction de liaison de l'ARNm optimale et que son remplacement par un bloc polyarginines n'entravait pas cette interaction contrairement à une mutation en polyalanines. La mutation conservative n'entraînant pas de défaut d'activité, il a été proposé que les motifs polylysines forment un groupe de charges positives qui pourrait interagir de manière non spécifique avec le squelette phosphaté d'un ARN. Ainsi, chez les Eucaryotes, l'ancrage de l'ARNm au niveau du motif en doigt de zinc serait facilité par ces interactions électrostatiques non spécifiques (Laurino *et al.*, 1998).

#### 4) Modèle de positionnement d'aIF2 $\beta$ sur aIF2 $\gamma$ .

La recherche d'homologues structuraux pour les deux domaines d'aIF2 $\beta$  a montré que le domaine N-terminal de la protéine archéenne présentait des similitudes avec diverses protéines impliquées dans la reconnaissance d'acides nucléiques, en particulier l'ADN (Gutierrez *et al.*, 2004). L'analyse de la position de ces protéines vis-à-vis de leur substrat ADN/ARN ainsi que les données biochimiques obtenues sur  $\beta$  ont conduit Gutierrez *et al.* (2004) à proposer un modèle de positionnement de la sous-unité  $\beta$  sur aIF2 $\gamma$  (**Figure 2.11**).



**Figure 2.11 :** Modèle de positionnement d'aIF2 $\beta$  de *M. thermoautotrophicum* sur la sous-unité aIF2 $\gamma$  de *P. abyssi* complexée à une molécule de GTP et à une molécule d'ARNt (d'après le positionnement réalisé par Schmitt *et al.*, 2002) (Gutierrez *et al.*, 2004). Dans ce modèle, les résidus de  $\beta$  identifiés comme nécessaires et suffisants à l'interaction avec  $\gamma$  se retrouvent bien en contact avec la grande sous-unité du facteur e/aIF2. D'autre part, le domaine N-terminal eucaryotique additionnel serait à proximité du site de liaison du GTP. Enfin, bien que  $\beta$  ait été impliqué dans la liaison à l'ARNm, sa taille exclut une interaction directe avec le polynucléotide. Dans ce modèle, le motif en doigt de zinc C-terminal pourrait plus probablement interagir avec la tige T de l'ARNt et par l'intermédiaire de cette interaction,  $\beta$  contrôlerait la validité de l'appariement codon-anticodon.

Ce modèle n'a cependant reçu aucune accréditation expérimentale jusqu'à présent. En particulier, une interaction entre e/aIF2 $\beta$  et l'ARNt n'a toujours pas été démontrée. Parmi les objectifs de mon travail de thèse, nous tenterons d'identifier la zone de  $\gamma$  en contact avec  $\beta$  et une analyse du rôle de  $\beta$  dans la liaison du Met-ARNt<sub>i</sub><sup>Met</sup> sera également entreprise.

### **5) Le domaine N-terminal d'eIF2 $\beta$ : une spécificité eucaryotique.**

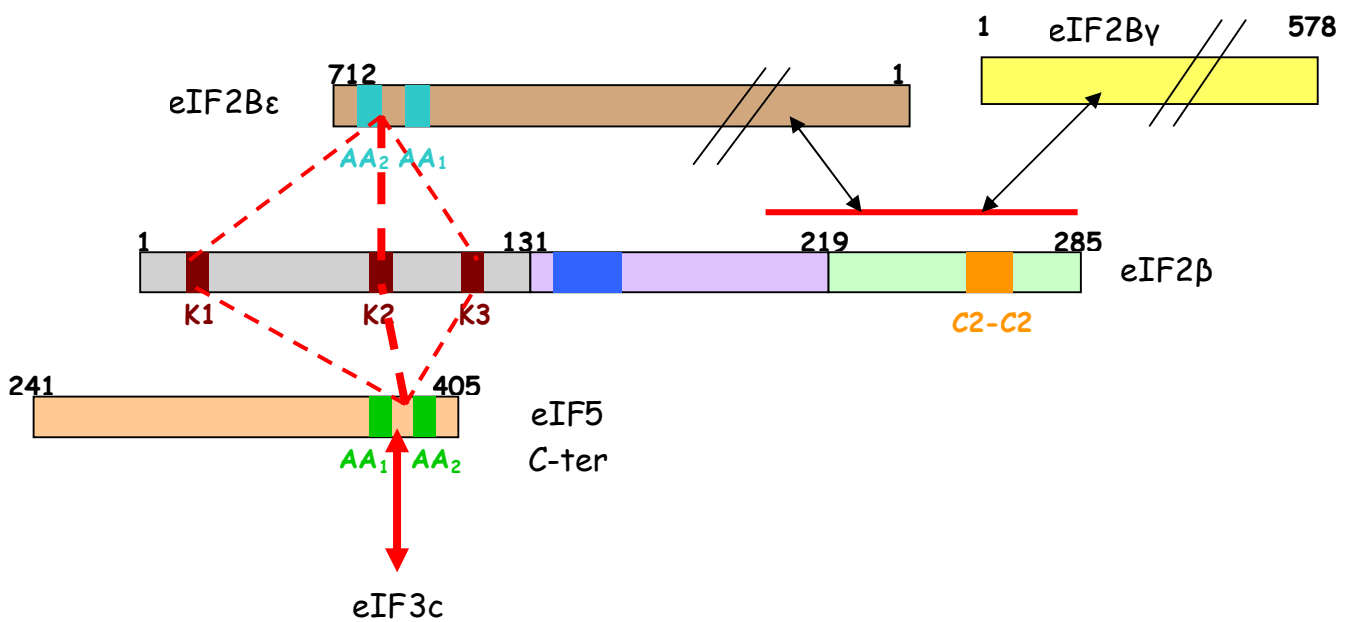
Chez les Eucaryotes, l'activité du facteur eIF2 est régulée par deux autres facteurs de démarrage : la GAP eIF5 et le facteur d'échange eIF2B. eIF2 $\beta$  se lie au facteur monomérique eIF5 ainsi qu'à la sous-unité catalytique eIF2B $\epsilon$  du facteur d'échange. On peut rappeler que le facteur eIF2B est composé de cinq sous-unités dont deux sous-unités catalytiques,  $\gamma$  et  $\epsilon$ , et trois sous-unités régulatrices,  $\alpha$ ,  $\beta$  et  $\delta$ . Ces sous-unités régulatrices forment un sous-complexe qui interagit avec la sous-unité  $\alpha$  du facteur eIF2 (voir Partie D de ce chapitre).

eIF2 $\beta$  lie eIF5 et eIF2B *via* son domaine N-terminal additionnel (voir **Figure 2.9**). Les motifs polylysines sont essentiels pour cette interaction (Asano *et al.*, 1999). Ainsi, une région de 22 acides aminés contenant le second motif polyK de  $\beta$  (résidus 68 à 89 chez l'homme) a d'abord été identifiée comme responsable de la liaison à eIF5 (Das *et al.*, 1997). En définitive, il semble que les trois blocs polyK aient une contribution additive sur la liaison à eIF5 : en effet, les mutants ne présentant que deux des trois motifs mutés en polyA conservent une capacité de liaison à eIF5 de 30 à 60% (Asano *et al.*, 1999).

L'interaction avec la sous-unité  $\epsilon$  d'eIF2B s'effectue également par l'intermédiaire de ces motifs polyK avec un rôle prépondérant joué par le second motif (K2). Dans ce cas, l'altération d'un seul motif polyK entraîne déjà un affaiblissement significatif de la liaison eIF2 $\beta$ -eIF2B $\epsilon$ . Il faut noter que ces deux interactions eIF2 $\beta$ -eIF5 et eIF2 $\beta$ -eIF2B $\epsilon$  sont mutuellement exclusives, ce qui est cohérent avec le fait qu'elles coordonnent des actions opposées sur le facteur eIF2 (hydrolyse du GTP dans le premier cas et échange GDP  $\rightarrow$  GTP dans le second cas). Un second site de liaison d'eIF2B sur  $\beta$  a été récemment mis en évidence : les sous-unités  $\delta$  et  $\epsilon$  du facteur d'échange interagiraient également avec le domaine C-terminal d'eIF2 $\beta$  (Kimball *et al.*, 1998).

Les domaines de liaison à eIF2 $\beta$  ont également été identifiés sur ces deux protéines. Chez tous les Eucaryotes, les protéines eIF5 et eIF2B $\epsilon$  présentent à leur extrémité C-terminale deux motifs très conservés riches en résidus à caractère acide et aromatique (Koonin, 1995). Ces deux motifs (12 résidus acides et/ou aromatiques pour le premier motif, 7 pour le second) sont généralement espacés d'une vingtaine de résidus beaucoup moins conservés. Au sein de chaque protéine, ces deux régions ont été impliquées dans la liaison à la sous-unité eIF2 $\beta$  (Asano *et al.*, 1999). Chez eIF2B $\epsilon$  la mutation en alanines des 7 résidus acides/aromatiques du second motif (mutant eIF2B $\epsilon$ -7A) entraîne un sévère défaut d'association à eIF2 $\beta$  *in vitro* et *in vivo* qui conduit à une perte d'activité du facteur eIF2B. La mutation des 12 résidus du premier motif (mutant eIF2B $\epsilon$ -12A) est létale chez la levure.

Au sein du facteur eIF5, ces deux motifs sont eux aussi impliqués dans la liaison d'eIF2 $\beta$  ainsi que dans la liaison de la sous-unité eIF3c. Dans le cas de l'association à eIF2 $\beta$ , une étude plus poussée des résidus conservés de ces motifs chez un eIF5 mammifère (rat) a montré que les doubles mutations E346A-E347A et E384A-E385A affaiblissaient la liaison eIF5-eIF2 $\beta$  sans modifier la force de la liaison à eIF3c. Cependant, l'altération d'un seul des deux motifs n'abolit pas entièrement la liaison à  $\beta$  et le facteur eIF5 peut ainsi assurer une fonction minimale *in vivo*. Au contraire, l'interaction d'eIF2 $\beta$  avec l'hexamutant E345A-E346A-E347A-E384A-E385A-E386A est négligeable et les cellules de levure privées du gène *TIF5* codant pour eIF5 et complémentées avec ce mutant ne sont plus viables, ce qui indique que la fonction *in vivo* de ce mutant est totalement abolie (Das and Maitra, 2000). Ainsi, les motifs riches en résidus acides et aromatiques du domaine C-terminal d'eIF5 interagissent directement avec les motifs polylysines du domaine N-terminal d'eIF2 $\beta$  et cette interaction est primordiale pour assurer le recrutement des deux facteurs au sein du MFC et garantir l'intégrité de leur fonction biologique (Asano *et al.*, 2001).



**Figure 2.12 :** Cartographie des interactions entre eIF2 $\beta$  et les facteurs de démarrage eIF5 et eIF2B. La numérotation relative aux protéines de levure a été utilisée dans ce schéma. Les sous-unités  $\gamma$  et  $\epsilon$  d'eIF2B n'ont pas été représentées dans leur totalité. De plus, les domaines respectifs de ces protéines qui interagissent avec la région C-terminale d'eIF2 $\beta$  n'ont pas encore été caractérisés à ce jour.

### **6) eIF5 : une GAP classique ?**

L'activité GTPasique de la sous-unité  $\gamma$  du facteur eIF2 nécessite, chez les Eucaryotes, une stimulation par le facteur eIF5. Cependant, contrairement aux GAPs classiques, eIF5 n'interagit pas directement avec sa GTPase  $\gamma$  mais avec la sous-unité eIF2 $\beta$  qui, elle-même,

est associée à  $\gamma$ . Au vu de cette originalité de partenariat, on peut se demander si le mécanisme général d'action d'eIF5 est similaire à celui des autres GAPs ou si, au contraire, il présente d'autres singularités.

Pour répondre à cette question, une analyse plus approfondie du facteur a été menée. eIF5 interagit avec  $\beta$  *via* son domaine C-terminal, mais les 100 premiers résidus d'eIF5 sont également nécessaires pour assurer une parfaite fonction GAP. En effet, un mutant d'eIF5 mammifère, eIF5 $\Delta$ 1-98, est toujours capable de lier eIF2 $\beta$  mais il n'est plus apte à déclencher l'hydrolyse du GTP sur le facteur eIF2 (Das *et al.*, 2001). Ce résultat suggère qu'eIF5 fonctionne comme une GAP classique, interagissant avec sa GTPase par l'intermédiaire de son domaine C-terminal et apportant son activité catalytique par son domaine N-terminal. La validation de cette hypothèse a été apportée en étudiant les résidus arginines et lysines du domaine N-terminal d'eIF5. La mutation systématique en alanine des Arg de ce domaine a montré que le mutant R15A est inactif *in vivo*. L'Arg15 est donc très probablement l'arginine catalytique de la GAP eIF5 qui intervient au sein d'un « Arginine finger ». De plus, elle est entourée par des résidus hydrophobes, Phe13 et Tyr14 qui pourraient permettre l'ancrage de la boucle catalytique de la GTPase au sein du coeur hydrophobe de sa GAP. Chez les autres GAPs, un ou plusieurs autres résidus Arg ou Lys interviennent de manière secondaire pour stabiliser la position de cette première arginine. Dans le cas d'eIF5, les résidus Lys33 et Lys55 semblent jouer ce rôle car leur mutation en alanine abolit la croissance des souches de levures  $\Delta$ TIF5 complémentées avec ces mutants. Tous ces mutants sont par contre parfaitement efficaces du point de vue de la liaison à eIF2 $\beta$  : leur rôle est donc bien un rôle catalytique dans l'activation de l'hydrolyse du GTP par eIF2.

Ainsi, eIF5 semble présenter un mécanisme classique d'activation de l'hydrolyse du GTP. Cependant, en l'absence du ribosome, eIF5 seul ne peut promouvoir l'hydrolyse du GTP porté par eIF2 au sein du complexe ternaire (Das *et al.*, 2001). En cela, la sous-unité 40S pourrait constituer une seconde GAP d'eIF2. Au cours de la phase d'élongation, l'activité GTPasique du facteur EF1A est stimulée en présence du ribosome (Dey *et al.*, 1995 ; Mohr *et al.*, 2002). Dans le cas d'EF1A, c'est le motif L7/L12 de la 50S qui stimule l'hydrolyse du GTP. Ce motif pourrait notamment induire un changement conformationnel d'EF1A qui placerait son domaine G dans une conformation catalytiquement active. Dans le cas d'eIF2, la sous-unité 60S n'est pas présente lors de l'hydrolyse du GTP puisqu'eIF2 est relargué de la 40S avant association des sous-unités ribosomales. Cependant, la petite sous-unité ribosomale pourrait elle aussi jouer un rôle dans l'activation conformationnelle d'eIF2, d'autant qu'eIF5 et eIF2 $\gamma$  n'interagissent pas directement. Enfin, certains résultats suggèrent qu'eIF5 et eIF2 s'associeraient préalablement à la formation du MFC. Les interactions de ces deux facteurs à eIF3 stimuleraient alors leur recrutement au sein du complexe de prédémarrage (Singh *et al.*,

2004). Dans ce cas, un moyen de contrôle de l'hydrolyse du GTP doit exister pour empêcher une dissociation prématurée du complexe de préinitiation sur un codon de démarrage erroné. Ce rôle pourrait être joué par eIF1 qui a effectivement un effet inhibiteur sur la fonction GAP d'eIF5 (Unbehaun *et al.*, 2004).

Chez les Archées, aucun homologue d'eIF5 n'est rencontré. Le facteur pourrait avoir une activité GTPasique intrinsèque plus élevée que celle de son homologue eucaryotique. En effet, au sein de la sous-unité eIF2 $\beta$ , une mutation permettant le démarrage sur un codon UUG avait été isolée : le mutant L254P. Ce dernier conférait à eIF2 une activité GTPasique intrinsèque, indépendante d'eIF5, et l'hydrolyse prématurée du GTP conduisait à un démarrage sur le codon mésapparié UUG. Chez aIF2 $\beta$ , un résidu proline est quasi-systématiquement rencontré au niveau de la position équivalente au résidu eucaryotique L254. aIF2 $\beta$ , et par conséquent le facteur aIF2, pourrait avoir une activité GTPasique plus grande qu'eIF2 (Thompson *et al.*, 2000). Dans ce cas, seul le ribosome serait nécessaire en tant que GAP et il pourrait éventuellement fournir l'arginine catalytique. La question de savoir comment l'hydrolyse du GTP est déclenchée chez les Archées reste encore ouverte et il n'est pas exclu qu'un autre facteur, non identifié à l'heure actuelle et non homologue à eIF5, joue le rôle de GAP dans ce mécanisme.

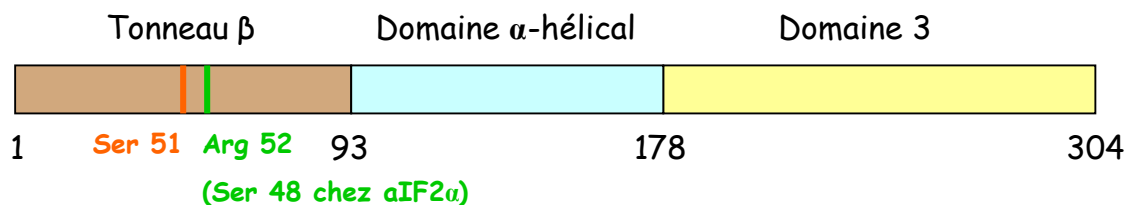
## **Partie D : La sous-unité $\alpha$ du facteur e/aIF2.**

### **1) Généralités.**

La sous-unité e/aIF2 $\alpha$ , codée par le gène *SUI2* (Cigan *et al.*, 1989), est une protéine d'environ 36 kDa (voir alignement de séquences en **Annexe 3**). Tout comme  $\beta$ , la sous-unité  $\alpha$  n'interagit qu'avec le noyau central de l'hétérotrimère, à savoir la sous-unité  $\gamma$  (Schmitt *et al.*, 2002). La zone d'interaction entre  $\alpha$  et  $\gamma$  était encore inconnue avant ce travail de thèse.

Chez les Eucaryotes, la sous-unité  $\alpha$  a un rôle clairement identifié dans la régulation du démarrage de la traduction. En effet, eIF2 $\alpha$  possède un résidu sérine universellement conservé dans son domaine N-terminal (Ser51 chez l'homme et chez la levure). Cette sérine est la cible de kinases spécifiques (PKR, HCR, GCN2 ou PERK) qui la phosphorylent en réponse à diverses conditions de stress. Sous forme phosphorylée, le facteur eIF2 devient alors un inhibiteur compétitif du facteur d'échange eIF2B. Ce dernier reste bloqué sous forme d'un complexe avec eIF2 phosphorylé et il ne peut donc plus recycler le facteur eIF2-GDP natif. La synthèse protéique est ainsi inhibée. Une interaction directe entre eIF2 $\alpha$  et les sous-unités régulatrices  $\alpha$ ,  $\beta$  et  $\delta$  du facteur d'échange a d'ailleurs été mise en évidence (Krishnamoorthy *et al.*, 2001).

Chez les Archées, le facteur eIF2B est absent. De plus, la Ser51 n'est pas universellement retrouvée dans les séquences d'aIF2 $\alpha$ . Ces observations ont donc conduit à l'hypothèse que ce rôle régulateur d' $\alpha$  n'existait pas dans le cas archéen. Cependant, une étude récente a montrée que le résidu Ser48 d'aIF2 $\alpha$  de *Pyrococcus horikoshii* pouvait être phosphorylé *in vitro* par la kinase PKR humaine (hPKR) ainsi que par une protéine de *P. horikoshii*, conséquemment identifiée comme étant l'homologue archéen de hPKR (Tahara *et al.*, 2004). Le résidu Ser48 (équivalent Arg52 chez eIF2 $\alpha$  de levure) est d'ailleurs très conservé dans toutes les séquences d'aIF2 $\alpha$ . Ainsi, chez certaines Archées, un rôle régulateur d'aIF2 $\alpha$  dans le démarrage de la traduction pourrait également être envisagé. Dans tous les cas, la fonction remplie par  $\alpha$  chez les Archées reste à préciser.

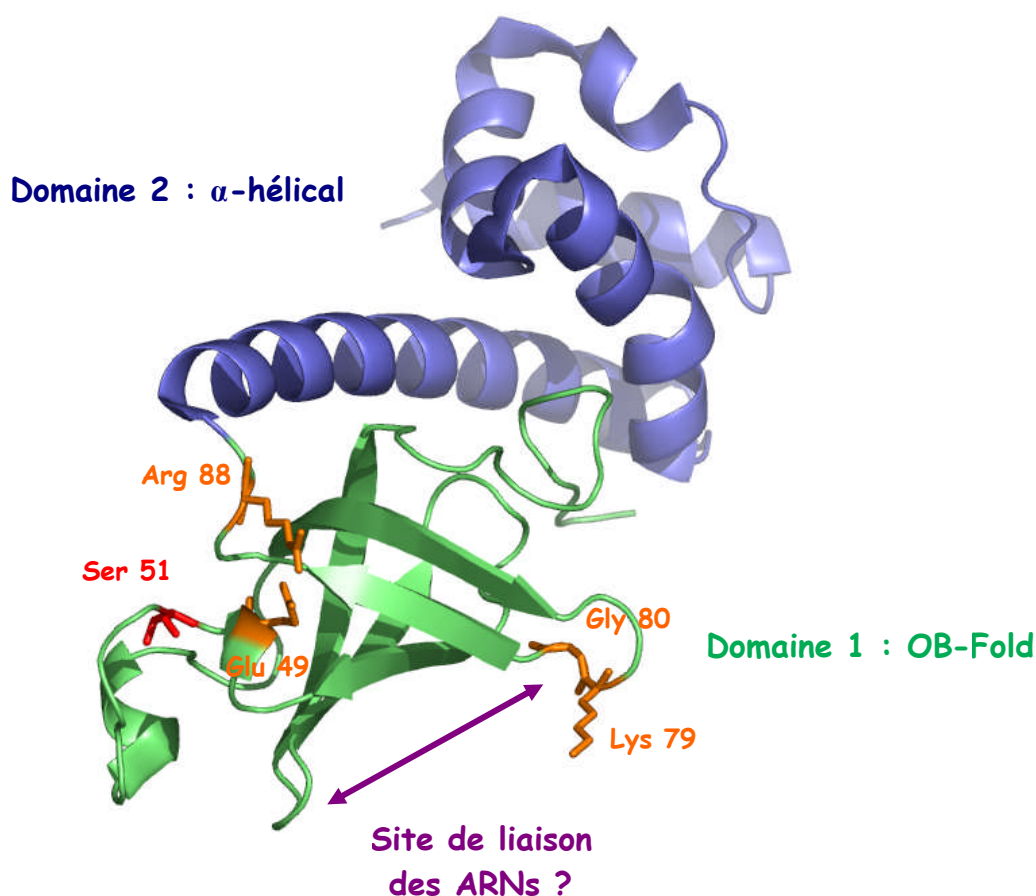


**Figure 2.13 :** Organisation structurale d'e/aIF2 $\alpha$  (Numérotation *S. cerevisiae*).

Chez les Eucaryotes, d'autres fonctions pour eIF2 $\alpha$  ont été proposées. Ainsi, la purification de complexes  $\alpha\gamma$  parfaitement actifs dans la formation du complexe ternaire suggérait qu' $\alpha$  participe à la liaison de l'ARNt initiateur. La sous-unité  $\alpha$  pourrait également lier les acides nucléiques. En particulier, des expériences de cross-linking ont soulevé la possibilité d'une interaction entre eIF2 $\alpha$  et l'ARN ribosomal 18S (Westermann *et al.*, 1980). Malgré ces quelques pistes, l'implication d'e/aIF2 $\alpha$  dans une fonction autre que celle de régulateur de la traduction eucaryotique reste à définir précisément. Dans ce contexte, l'analyse d'un rôle éventuel de la sous-unité  $\alpha$  dans la liaison du Met-ARNt $_i^{\text{Met}}$  et dans la fonction générale du facteur aIF2 a constitué un objectif majeur de mon travail de thèse.

## 2) Structure de la sous-unité $\alpha$ .

La sous-unité  $\alpha$  est composée de trois domaines (**Figure 2.13**). Pendant mon travail de thèse, la structure des deux domaines N-terminaux a été résolue dans le cas des protéines humaine et de levure (Nonato *et al.*, 2002 ; Dhaliwhal and Hoffman, 2003). La **Figure 2.14** représente la protéine eIF2 $\alpha$  humaine.



**Figure 2.14 :** Structure des deux domaines N-terminaux d'eIF2 $\alpha$  de levure à 2.86 Å de résolution (Dhaliwal and Hoffman, 2003).

Le domaine N-terminal d'eIF2 $\alpha$  (résidus 1 à 87) adopte un repliement de type OB-fold fréquemment rencontré dans les protéines liant les acides nucléiques. Une homologie structurale a ainsi été mise en évidence avec le facteur IF1/eIF1A, les protéines ribosomales S1 et S17, la polynucléotide phosphorylase d'*E. coli* et le domaine N-terminal de l'AspRS. Le domaine 1 d'eIF2 $\alpha$  pourrait être impliqué dans l'interaction avec un partenaire de type ARN, notamment l'ARNm, puisque des mutations au sein de ce domaine altèrent la sélection du bon codon de démarrage (mutants P13S et V19F chez l'homme). Le second domaine de cette structure (résidus 88 à 182) est exclusivement hélical.

Un autre objectif de mon travail de thèse a été la détermination de la structure de la sous-unité entière, libre et en complexe avec la sous-unité  $\gamma$ .

### **3) Le recyclage GDP $\rightarrow$ GTP chez les Eucaryotes.**

Le facteur d'échange eIF2B est composé de 5 sous-unités :  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  et  $\epsilon$ , respectivement codées par les gènes *GCN3*, *GCD7*, *GCD2*, *GCD1* et *GCD6*. Ce multimère forme deux sous-complexes : un sous-complexe catalytique (sous-unités  $\gamma$  et  $\epsilon$ ) et un sous-complexe régulateur (sous-unités  $\alpha$ ,  $\beta$  et  $\delta$ ) (Pavitt *et al.*, 1998). Le mécanisme par lequel eIF2B catalyse le recyclage d'eIF2 en fin de démarrage de la traduction est encore méconnu. De même, l'effet inhibiteur de la phosphorylation d'eIF2 $\alpha$  sur l'activité catalytique du facteur d'échange est encore largement incompris. Le facteur eIF2B présente une affinité bien supérieure pour le facteur eIF2 phosphorylé en comparaison avec le substrat natif (Pavitt *et al.*, 1998).

La sous-unité eIF2B $\alpha$  est directement impliquée dans l'inspection de l'état de phosphorylation d'eIF2 $\alpha$ . En effet, en présence d'eIF2 $\alpha$  phosphorylé au niveau de la sérine 51, eIF2-( $\alpha$ P), un facteur eIF2B privé de sa sous-unité  $\alpha$  retrouve une activité d'échange nucléotidique intacte, surmontant ainsi l'effet inhibiteur dû à la phosphorylation d'eIF2 $\alpha$  (Kimball *et al.*, 1998).

Les sous-unités  $\beta$  et  $\delta$  sont également impliquées dans cette inspection. Différents mutants de ces deux sous-unités permettant de surmonter l'effet inhibiteur de la phosphorylation d'eIF2 $\alpha$  ont pu être caractérisés (Kimball *et al.*, 1998) ; Pavitt *et al.*, 1997). Ces mutations n'aboliraient pas l'affinité préférentielle d'eIF2B pour eIF2-( $\alpha$ P). Par contre, elles permettraient au facteur eIF2B de reconnaître le facteur eIF2 phosphorylé comme un substrat actif pour la réaction d'échange nucléotidique (Pavitt *et al.*, 1998).

Ainsi, le sous-complexe eIF2B $\alpha\beta\delta$  ne possède pas d'activité d'échange nucléotidique. Par contre, il est susceptible de retenir la forme phosphorylée d'eIF2, piégeant ainsi le facteur eIF2B au sein d'un complexe inactif. Une interaction directe d'eIF2 $\alpha$  avec ce sous-complexe a d'ailleurs été mise en évidence et elle implique entre autres les résidus Glu49, Lys79, Gly80 et Arg88 (**Figure 2.14**). Ces résidus correspondent à deux zones du domaine 1 relativement éloignées, ce qui suggère que deux points de contact distincts doivent exister entre eIF2 $\alpha$  et le sous-complexe régulateur d'eIF2B. Le site de fixation de différentes kinases d'eIF2 $\alpha$ , en particulier PKR, semble également correspondre à cette zone de la protéine. En particulier, les résidus 79 à 83 sont impliqués dans la reconnaissance du substrat  $\alpha$  par sa kinase (Dhaliwal and Hoffman, 2003). D'autre part, plusieurs mutants du domaine N-terminal d'eIF2 $\alpha$  ont permis de supprimer l'effet inhibiteur de sa phosphorylation. Ces mutations impliquent les résidus Ile58, Leu84, Arg88 et Val89. Ils pourraient ainsi être directement responsables d'une interaction avec le noyau régulateur d'eIF2B, en particulier la sous-unité  $\alpha$  (Vasquez de Aldana *et al.*, 1993).

L'activité catalytique d'eIF2B dans la réaction d'échange est assurée par les deux sous-unités  $\gamma$  et  $\epsilon$ . Ce sous-complexe isolé est en effet capable de catalyser la réaction d'échange et il n'est pas sensible à l'état de phosphorylation d'eIF2. eIF2B $\epsilon$  se lie à son substrat par l'intermédiaire d'eIF2 $\beta$ . C'est la sous-unité catalytique du facteur et eIF2B $\gamma$ , en formant un sous-complexe avec eIF2B $\epsilon$ , permet de stabiliser l'interaction GEF-GTPase (Pavitt *et al.*, 1998). Ainsi, un modèle d'action du facteur eIF2B, décomposé en deux temps, a été proposé (Pavitt *et al.*, 1998). Les trois sous-unités régulatrices d'eIF2B formeraient une première interaction, non productive, avec eIF2 par l'intermédiaire de sa sous-unité  $\alpha$ . Cette interaction ne permettrait pas de placer les sous-unités catalytiques  $\gamma$  et  $\epsilon$  dans un contexte favorable pour lier la GTPase et assurer le recyclage GDP  $\rightarrow$  GTP. En l'absence de phosphorylation d'eIF2 $\alpha$ , le facteur eIF2B pourrait alors subir une isomérisation qui permettrait de placer les sous-unités catalytiques  $\gamma$  et  $\epsilon$  à proximité d'eIF2 $\gamma$ . Cette seconde interaction, productive, déclencherait alors la réaction d'échange. La phosphorylation d'eIF2 $\alpha$  empêcherait ainsi l'isomérisation du facteur et celui-ci resterait bloqué dans une interaction non productive avec son substrat. La sous-unité eIF2 $\alpha$ , bien que ne participant pas directement à la réaction d'échange, contribuerait cependant à l'énergie de liaison du complexe eIF2-GDP-eIF2B en permettant la formation d'un premier complexe non productif. D'ailleurs, en son absence, le Km d'eIF2-GDP dans la réaction d'échange est augmenté d'un ordre de grandeur (Nika *et al.*, 2001).

La question du mécanisme exact par lequel survient la réaction d'échange reste toutefois ouverte, le modèle décrit ci-dessus n'ayant pas encore reçu de validation expérimentale. De plus, en l'absence d'homologues du sous-complexe catalytique d'eIF2B

chez les Archées, le problème du recyclage d'aIF2-GDP reste entier. En effet, seules les sous-unités  $\alpha$ ,  $\beta$  et  $\delta$  semblent représentées au sein de ces organismes. Cependant, ces orthologues archéens des sous-unités régulatrices d'eIF2B pourraient ne pas avoir de lien fonctionnel avec le démarrage de la traduction. Ainsi, chez les Archées, le recyclage GDP  $\rightarrow$  GTP pourrait survenir de manière spontanée (Kyripides and Woese, 1998).

## **OBJECTIFS.**

Le facteur e/aIF2 joue un rôle crucial dans le démarrage de la traduction. Cependant, comme nous l'avons vu tout au long de ces deux chapitres d'introduction, beaucoup de questions concernant son mode de fonctionnement restent encore ouvertes. Le travail réalisé au cours de cette thèse avait pour objectif de poursuivre l'étude de ce facteur d'un point de vue fonctionnel mais également structural. Un aspect majeur de la fonction d'e/aIF2 est au centre de cette étude : la reconnaissance spécifique de l'ARNt initiateur par le facteur hétérotrimérique. En effet, e/aIF2 $\gamma$  est homologue au facteur d'élongation qui reconnaît indistinctement tous les ARNts cellulaires. Pourtant, e/aIF2 ne sélectionne que le Met-ARNt<sub>i</sub><sup>Met</sup>. L'une des questions fondamentales qui reste posée sur le facteur de démarrage est donc de savoir par quel mécanisme il assure cette reconnaissance spécifique : quels sont les déterminants reconnus sur l'ARNt ? Au sein du facteur e/aIF2, quelles sont les sous-unités impliquées dans la liaison au polynucléotide ? En effet, si un rôle de  $\gamma$  semble probable du fait de son homologie avec EF1A, aucun rôle clair pour  $\beta$  ou pour  $\alpha$  n'a pu être défini dans cette interaction.

Cette première question en soulève une autre, plus générale : quelle est la concertation qui existe entre les trois sous-unités du facteur e/aIF2 ? En effet, contrairement à son pendant bactérien IF2 et à l'homologue de  $\gamma$ , EF1A, e/aIF2 dispose de trois sous-unités pour remplir sa fonction. Il est donc clair que leur présence simultanée et leur action coordonnée sont un point essentiel pour la fonction biologique du facteur. Une étude de l'agencement structural de ce facteur est primordiale pour comprendre la communication qui s'effectue au sein de l'hétérotrimère. Avant ce travail de thèse, il manquait encore des données structurales sur le domaine 3 d'e/aIF2 $\alpha$  ainsi que sur le domaine N-terminal additionnel d'eIF2 $\beta$ . D'autre part,

aucune information n'était disponible sur l'organisation structurale du facteur et sur les zones des protéines  $\alpha$  et  $\gamma$  impliquées dans ce réseau d'interactions.

Actuellement, de nombreuses études structurales portent sur les protéines d'organismes extrémophiles, en particulier les Archées. Ces choix se justifient par une plus grande stabilité des macromolécules étudiées. Le génome de nouvelles Archées, *Pyrococcus abyssi* et *Sulfolobus solfataricus*, ayant été récemment séquencé (Lecompte *et al.*, 2001 ; She *et al.*, 2001), nous avons porté notre choix d'étude du facteur aIF2 sur ces deux organismes, dans le but de modéliser la fonction du facteur eucaryotique, tout en gardant à l'esprit les différences qui existent probablement entre les deux mécanismes de démarrage. A mon arrivée au laboratoire, les trois sous-unités du facteur de *P. abyssi* étaient disponibles et la structure de la sous-unité  $\gamma$  résolue. C'est donc sur le système de *P. abyssi* que nous avons démarré nos travaux.

Dans les chapitres qui suivent, les protéines des archées *P. abyssi* et *S. solfataricus* seront désignées par les notations respectives Pa-aIF2 et Ss-aIF2. La notation aIF2 désignera le facteur hétérotrimérique aIF2 en général.

*Chapitre III :*  
*Cartographie*  
*Fonctionnelle Du Facteur*  
*d'Initiation aIF2*



## **Partie A : Reconnaissance spécifique de l'ARNt initiateur par le facteur aIF2.**

Le facteur d'initiation e/aIF2 intervient de manière hautement spécifique au cours du démarrage de la synthèse protéique puisqu'il réalise une discrimination positive de l'ARNt initiateur méthionylé face aux autres ARNs de transfert présents dans la cellule. Au contraire, le facteur d'élongation e/aEF1-A est capable de reconnaître tous les autres ARNs et de les acheminer sur le ribosome. Pourtant, ces différents ARNs présentent une structure tridimensionnelle globalement identique. Ce sont donc des caractéristiques beaucoup plus fines qui doivent être à l'origine de cette reconnaissance spécifique du Met-ARNt initiateur par e/aIF2.

Nous nous sommes donc attachés dans un premier temps à identifier les déterminants présents sur l'ARNt initiateur qui expliqueraient sa prise en charge par le seul facteur d'initiation 2. Pour cela, nous avons mis en place un test permettant de mesurer l'affinité de l'hétérotrimère e/aIF2 pour différents ARNs. Notre étude s'est portée sur le facteur d'initiation aIF2 de l'archée *Pyrococcus abyssi*, noté par la suite Pa-aIF2. Le **Tableau 3.1** reprend les caractéristiques des différentes protéines de l'hétérotrimère.

<b>Protéine</b>	<b>Poids moléculaire</b>	<b>Point isoélectrique</b>	<b>Facteur d'absorption molaire en L.cm<sup>-1</sup>.mol<sup>-1</sup></b>
Pa-aIF2 $\alpha$	31.9 kDa	8.72	$39.7 \times 10^3$
Pa-aIF2 $\beta$	16.6 kDa	8.07	$9.5 \times 10^3$
Pa-aIF2 $\gamma$	44.9 kDa	7.99	$2.6 \times 10^3$

**Tableau 3.1:** Caractéristiques générales des différentes sous-unités du facteur aIF2 de *P. abyssi* déterminées à l'aide du programme *DNAid* (Dardel et Bensoussan, 1988).

### **1) Clonage des gènes codant pour les 3 sous-unités du facteur aIF2 de *P. abyssi*.**

Dans la perspective de pouvoir reproduire les mesures de ces constantes d'affinité sur les assemblages hétérodimériques et sur les monomères de Pa-aIF2, il nous fallait un système d'expression qui permette la production séparée des trois sous-unités d'aIF2 de *P. abyssi*. Les conditions de croissance de cette archée hyperthermophile ( $T_{\text{optimale}} = 103^{\circ}\text{C}$  ;  $P_{\text{optimale}} = 200$  bars) n'étant pas reproductibles au laboratoire, un mode de production chez la bactérie *Escherichia coli* a été envisagé. Le séquençage intégral du génome de l'archée *Pyrococcus*

*abyssi* (Lecompte et al., 2001) a permis d'identifier les gènes PAB0568, PAB0959 et PAB2040 codant respectivement pour les sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$  de Pa-aIF2. Préalablement à ce travail de thèse, chacun de ces gènes a été cloné au laboratoire dans une version modifiée des vecteurs d'expression procaryotiques pET3a et pET15b.

Les vecteurs pET3a et pET15b possèdent une origine de réplication de type pBR322, une séquence promotrice reconnue par l'ARN-polymérase du bactériophage T7, un site de clonage multiple et un terminateur reconnu par l'ARN-polymérase du bactériophage T7. Ils portent également le gène codant pour la protéine LacI et le gène codant pour la  $\beta$ -lactamase, protéine qui confère à la bactérie transformée une résistance à l'ampicilline. Dans le cas du pET15b, une séquence codant pour un hexapeptide histidine (6xHis-Tag) est également présente. Les versions modifiées de ces vecteurs construites au laboratoire comportent en fait un nombre de sites uniques utilisables pour le clonage plus important que le vecteur de départ. En effet, une séquence additionnelle comportant les sites de restriction indiqués ci-après a été ajoutée entre les sites *NdeI* et *BamHI* du site de clonage multiple.

*NdeI* / *SmaI* / *XhoI* / *KpnI* / *SacII* / *NotI* / *SalI* /  $\Delta$ *BamHI*

Les deux vecteurs ainsi construits ont été nommés pET3a $\alpha$ lp et pET15b $\alpha$ lp. Les gènes codant pour chacune des trois sous-unités de Pa-aIF2 ont été clonés entre les sites de restriction *NdeI* et *XhoI*, donnant naissance aux vecteurs d'expression pET3a $\alpha$ lp $\alpha$ , pET3a $\alpha$ lp $\beta$ , pET3a $\alpha$ lp $\gamma$ , pET15b $\alpha$ lp $\alpha$ , pET15b $\alpha$ lp $\beta$  et pET15b $\alpha$ lp $\gamma$  qui permettent d'exprimer séparément les 3 sous-unités du facteur d'initiation.

## **2) Purification des 3 sous-unités du facteur Pa-aIF2.**

### **1. Mode de surproduction adopté.**

Les vecteurs pET3a $\alpha$ lp $\alpha$ , pET3a $\alpha$ lp $\beta$  et pET3a $\alpha$ lp $\gamma$  ont été utilisés pour transformer les souches BL21(DE3) d'*Escherichia coli*. Les ADN d'archées comportent fréquemment certains codons tels que AGA, AGG (codant pour Arg) ou ATA (codant pour Ile) qui sont rarement rencontrés sur l'ADN d'*E. coli*. La bactérie ne produit que de faibles quantités des ARNts capables de lire ces codons. C'est pourquoi, il faut également lui permettre de produire les ARNts capables de déchiffrer ces codons rares éventuellement présents sur les gènes codant pour les trois sous-unités de Pa-aIF2. Pour cela, les cellules bactériennes sont co-transformées avec le plasmide portant le gène d'intérêt et un plasmide permettant l'expression de ces ARNts particuliers. Dans le cas des sous-unités  $\alpha$  et  $\beta$ , c'est le plasmide pSBETa (Schenk et al., 1995) qui a été utilisé. Ce plasmide confère une résistance à la kanamycine à la

bactérie transformée, permettant ainsi la sélection des seules cellules doublement transformées. Pour la production de la sous-unité  $\gamma$ , c'est le plasmide pSJS1240, qui confère une résistance à la spectinomycine, qui a été utilisé dans la co-transformation des cellules BL21.

Les gènes d'intérêt sont placés sous contrôle d'un promoteur reconnu par l'ARN-polymérase du bactériophage T7 au sein du vecteur d'expression pET3alpa. Le gène codant pour cette ARN-polymérase est lui-même présent sur l'ADN chromosomique des souches d'*E.coli* BL21(DE3) et il est placé sous contrôle du promoteur de l'opéron lactose, activé en présence d'IPTG. L'ajout d'IPTG 1 mM dans les cultures bactériennes en phase exponentielle de croissance permet donc d'enclencher la surproduction des protéines d'intérêt.

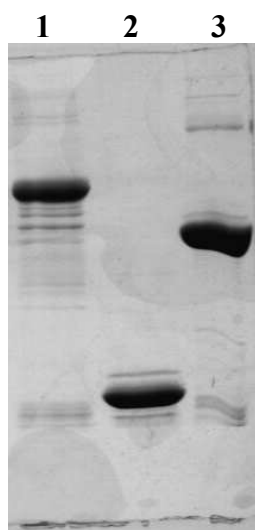
## **2. Purification de Pa-aIF2 $\alpha$ , Pa-aIF2 $\beta$ et Pa-aIF2 $\gamma$ .**

Le mécanisme détaillé de purification de ces protéines est repris dans la partie **Matériels et Méthodes**.

Brièvement, nous avons tiré parti du fait que les protéines à purifier proviennent de l'archée hyperthermophile *P. abyssi*. Ainsi, nous avons soumis l'extrait protéique obtenu après lyse des cellules à un chauffage à 80°C. A cette température, les protéines d'intérêt sont stables tandis que la majeure partie des protéines bactériennes, thermosensibles, précipite et est ainsi séparée par centrifugation. Un traitement au sulfate de streptomycine a ensuite été appliqué sur la solution contenant les protéines thermostables, pour précipiter les acides nucléiques. Puis les protéines Pa-aIF2 ont subi une précipitation au sulfate d'ammonium.

Malgré des points isoélectriques assez basiques (**Tableau 3.1**) pour les protéines  $\beta$  et  $\gamma$ , nous n'avons trouvé aucune colonne échangeuse d'ions capable de retenir spécifiquement la protéine d'intérêt dans des conditions qui préservent son intégrité. Trois passages au travers de colonnes échangeuses d'ions (anions et cations) permettent d'éliminer la plupart des protéines contaminantes et les acides nucléiques de l'extrait protéique. Un polissage final est effectué par tamisage moléculaire.

La sous-unité  $\alpha$ , au contraire, s'accroche sur une résine échangeuse de cations dans des conditions de pH et de force ionique qui maintiennent sa stabilité. Une résine échangeuse d'anions a d'abord été employée pour éliminer les acides nucléiques. Puis la sous-unité  $\alpha$  a été purifiée par chromatographie échangeuse de cations sur une colonne S-Sépharose. Enfin, un tamis moléculaire a permis de séparer les derniers contaminants.



**Figure 3.2 :** migration électrophorétique des 3 sous-unités purifiées d'aIF2 de *P. abyssii* sur un gel de polyacrylamide 12%. Piste 1 : aIF2 $\gamma$  (45 kDa) ; Piste 2 : aIF2 $\beta$  (16.6 kDa) ; Piste 3 : aIF2 $\alpha$  (32 kDa).

### 3) Mesure de l'affinité de Pa-aIF2 pour l'ARNt.

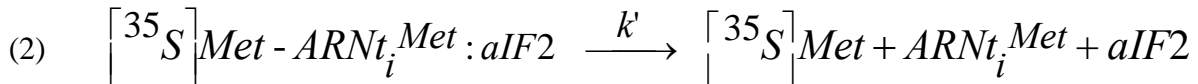
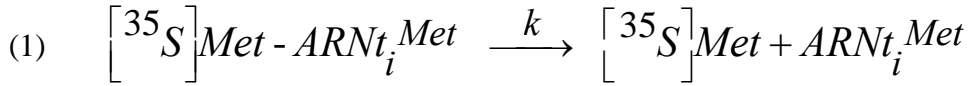
#### 1. Principe de la méthode.

La méthode utilisée pour mesurer les constantes d'affinité des complexes Pa-aIF2:ARNt repose sur la capacité du facteur Pa-aIF2 à protéger l'ARNt aminoacylé contre la déacylation spontanée qui survient dans le milieu réactionnel. Ce type de protection a déjà été observé pour le facteur d'élongation EF1A (Pingoud *et al.*, 1977). En effet, au sein du complexe ternaire EF1A:GTP:aminoacyl-ARNt, la liaison ester qui lie l'acide aminé à son ARNt est stabilisée car moins exposée au solvant et par conséquent, son hydrolyse spontanée est défavorisée. Ceci est d'autant plus vrai que la quantité d'EF1A présente dans le milieu, et par conséquent la quantité de complexe ternaire formé, est grande. Du fait des rôles symétriques joués par EF1A et Pa-aIF2 respectivement dans l'élongation et le démarrage de la traduction, nous avons supposé que Pa-aIF2 pouvait jouer le même rôle protecteur qu'EF1A vis-à-vis du Met-ARNt<sub>i</sub><sup>Met</sup>. Cette hypothèse semblait d'autant plus raisonnable qu'EF1A et Pa-aIF2 $\gamma$  ont des structures tridimensionnelles homologues et par conséquent, un mode de reconnaissance de l'ARNt très probablement similaire.

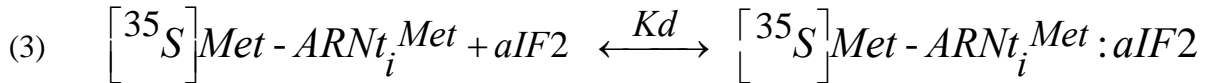
Nous avons donc cherché à suivre la cinétique de déacylation du Met-ARNt<sub>i</sub><sup>Met</sup> en présence de concentrations croissantes du facteur Pa-aIF2. Pour pouvoir mesurer les paramètres cinétiques de la déacylation, nous avons utilisé un ARNt marqué radioactivement avec de la [<sup>35</sup>S]-Met. Pour éviter l'hydrolyse du GTP en GDP au cours de la mesure, nous avons utilisé un analogue non hydrolysable du GTP, le GDPNP. D'autre part, nous avons

supposé que l'ARNt non aminoacylé n'interagissait pas avec le facteur Pa-aIF2 (cette hypothèse a d'ailleurs été vérifiée expérimentalement comme nous le verrons par la suite).

En considérant ces différentes hypothèses, les réactions de déacylation qui vont avoir lieu dans le milieu sont donc les suivantes :



D'autre part, l'équilibre suivant est réalisé :



Cet équilibre étant atteint quasi instantanément, à tout instant t on peut écrire :

$$(4) \quad Kd = \frac{[\text{Met} - \text{ARNt}]_t \times [\text{aIF2}]_t}{[\text{Met} - \text{ARNt} : \text{aIF2}]_t}$$

La vitesse de déacylation, qui prend en compte toutes les espèces aminoacylées, s'écrit alors comme suit :

$$(5) \quad v = - \left( \frac{d[\text{Met} - \text{ARNt}]_t}{dt} + \frac{d[\text{Met} - \text{ARNt} : \text{aIF2}]_t}{dt} \right) = \frac{d[\text{ARNt}]_t}{dt}$$

$$= k[\text{Met} - \text{ARNt}]_t + k'[\text{Met} - \text{ARNt} : \text{aIF2}]_t$$

$$= \left( k + \frac{k' [\text{aIF2}]_t}{Kd} \right) [\text{Met} - \text{ARN}]_t \quad \text{d'après la relation (4)}$$

La conservation des espèces donne la relation (6) suivante :

$$\begin{aligned} [Met - ARNt]_0 &= [Met - ARNt]_t + [ARNt]_t + [Met - ARNt : aIF2]_t \\ &= \left( 1 + \frac{[aIF2]_t}{Kd} \right) [Met - ARNt]_t + [ARNt]_t \quad \text{d'après (4)} \end{aligned}$$

Cette dernière équation permet d'exprimer  $[Met-ARNt]_t$  en fonction d' $[ARNt]_t$  dans la relation (5) qui devient :

$$\frac{d[ARNt]_t}{dt} = \left( \frac{k \cdot Kd + k' [aIF2]_t}{Kd + [aIF2]_t} \right) ([Met - ARNt]_0 - [ARNt]_t)$$

Si l'on se place dans des conditions telles que la quantité d'aIF2 engagée dans la formation du complexe ternaire est très faible par rapport à sa quantité initiale (il faut donc se placer dans des conditions expérimentales telles que  $[aIF2]_0 \gg [Met-ARNt]_0$ ), alors  $[aIF2]_t \approx [aIF2]_0$  et (6) devient :

$$\frac{d[ARNt]_t}{dt} = B \times ([Met - ARNt]_0 - [ARNt]_t) \quad \text{avec} \quad B = \frac{(k - k') \cdot Kd}{Kd + [aIF2]_0} + k'$$

L'intégration de cette équation avec la condition limite  $[ARNt]_t \rightarrow 0$  quand  $t \rightarrow 0$  (car on suppose qu'à l'instant  $t = 0$ , il n'y a pas d'ARNt déacylé) nous donne :

$$(8) \quad [Met - ARNt]_0 - [ARNt]_t = e^{-Bt} \quad \text{où} \quad B = \frac{(k - k') \cdot Kd}{Kd + [aIF2]_0} + k'$$

Ainsi, pour chaque concentration en Pa-aIF2, la cinétique de déacylation du Met-ARNt est suivie en mesurant, à différents temps, la quantité d'ARNt aminoacylé encore présent dans le mélange réactionnel. L'équation (8) modélise l'évolution de cette quantité. Le programme MC-Fit (Dardel, 1994) permet alors de déduire la constante de temps  $\tau$  relative à cette exponentielle et par conséquent la constante de vitesse  $B = 1/\tau$ .

Le tracé de la courbe  $B=f([Pa-aIF2]_0)$  permet alors d'obtenir la valeur du  $Kd$  de l'assemblage Pa-aIF2:ARNt considéré grâce au programme MC-Fit (Dardel, 1994).

Lorsque le  $K_d$  n'est pas grand devant la concentration d'ARNt, on ne peut plus supposer  $[aIF2]_t \approx [aIF2]_0$ . On utilise alors un traitement qui permet de corriger la concentration  $[aIF2]_t$  compte tenu de la fraction d'aIF2 liée à l'ARNt.

## **2. Mise en œuvre expérimentale.**

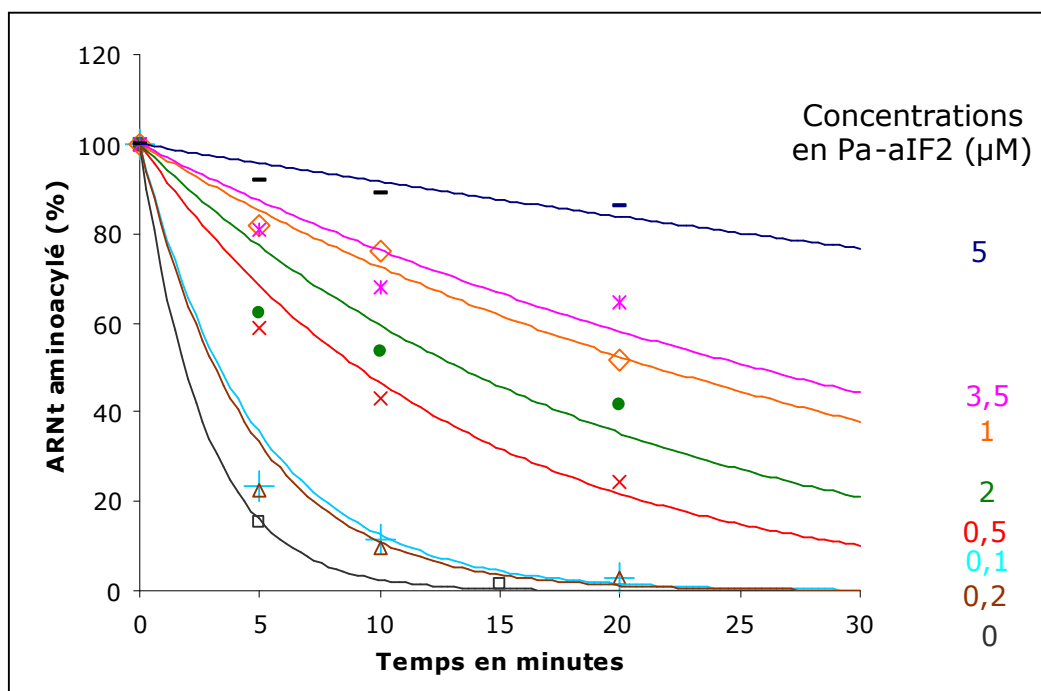
De manière pratique, pour pouvoir suivre la déacylation de l'ARNt sur un temps raisonnable, nous nous sommes placés dans un milieu suffisamment déacylant pour permettre une hydrolyse quasi complète de la liaison ester du Met-ARNt en 30 minutes. Pour cela, le milieu réactionnel choisi est un tampon 20 mM HEPES pH 8.0, 100 mM KCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mM EDTA, 5% glycérol, 0.2 mg/mL BSA. La force ionique du milieu est un paramètre important. En effet, des essais à plus haute force ionique (jusqu'à 500 mM KCl) ont montré que le sel défavorise la protection du Met-ARNt par Pa-aIF2, augmentant ainsi le  $K_d$  mesuré. La température de mesure de ces cinétiques a été fixée à 51°C. L'hétérotrimère Pa-aIF2 étant issu d'une archée hyperthermophile, une température assez élevée doit être utilisée pour avoir une bonne activité de protection du facteur. Des tests réalisés à plus basse température ont montré une protection de l'ARNt par Pa-aIF2 beaucoup plus faible pour une même concentration protéique. A contrario, la majeure partie de ces mesures ayant été réalisées avec l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli*, un test à une température plus proche de la température optimale des protéines de *P. abyssi* (80°C) n'a pas été envisagé, par crainte d'une dégradation thermique du polynucléotide.

Le protocole de mesure des constantes de vitesse  $B$  des réactions de cinétique pour chaque concentration en Pa-aIF2 est détaillé dans la partie **Matériels et Méthodes**. La **Figure 3.3** donne un exemple des courbes de déacylation expérimentales obtenues et la **Figure 3.4** illustre le tracé de la courbe  $B=f([Pa-aIF2]_0)$  de laquelle est déduit le  $K_d$  du complexe étudié.

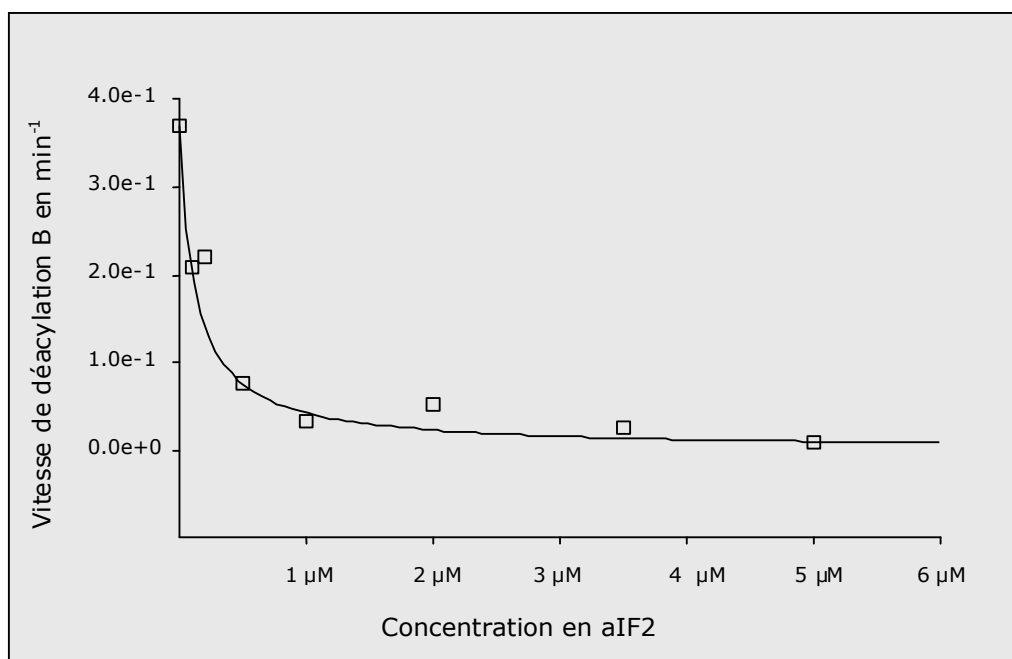
## **3. Résultats préliminaires et mesure du $K_d$ pour le complexe Pa-aIF2:Met-ARNt<sub>f</sub><sup>Met</sup>.**

Dans un premier temps, nous nous sommes donc intéressés au complexe Pa-aIF2:Met-ARNt<sub>f</sub><sup>Met</sup> de *P. abyssi*. L'ARNt initiateur de *P. abyssi* a été purifié puis aminoacylé avec de la [<sup>35</sup>S]-Met (10000 dpm/pmole) comme décrit dans la partie **Matériels et Méthodes**. L'hétérotrimère Pa-aIF2 a été obtenu en mélangeant dans un rapport équimolaire les trois sous-unités purifiées de Pa-aIF2. L'ajustement de la stœchiométrie de ces complexes a été réalisé grâce à la détermination expérimentale des  $\epsilon_{\text{molaires}}$ , celle-ci étant rendue possible par la mesure du contenu en zinc de  $\beta$  et  $\gamma$  à l'aide d'un spectromètre de flamme. La stœchiométrie

de l'assemblage a finalement été vérifiée par dépôt d'un aliquot sur gel d'électrophorèse natif. Un exemple de gel est illustré pour l'hétérodimère Pa-aIF2 $\alpha\gamma$  sur la Figure 3.9.



**Figure 3.3 :** Mesure de la constante de dissociation du Met-ARNt<sub>f</sub><sup>Met</sup> d'*E.coli* pour l'hétérotrimère aIF2 de *P. abyssi*. Cinétique de déacylation du Met-ARNt<sub>f</sub><sup>Met</sup> en présence de concentrations croissantes du facteur Pa-aIF2. Chaque courbe de déacylation est modélisée pour une exponentielle simple.



**Figure 3.4 :** Mesure de la constante de dissociation du Met-ARNt<sub>f</sub><sup>Met</sup> d'*E.coli* pour l'hétérotrimère Pa-aIF2 de *P. abyssi*. Tracé des vitesses de déacylation en fonction de la concentration initiale en aIF2. Le programme MC-Fit permet de déduire de cette courbe la valeur du K<sub>d</sub> pour le complexe considéré.

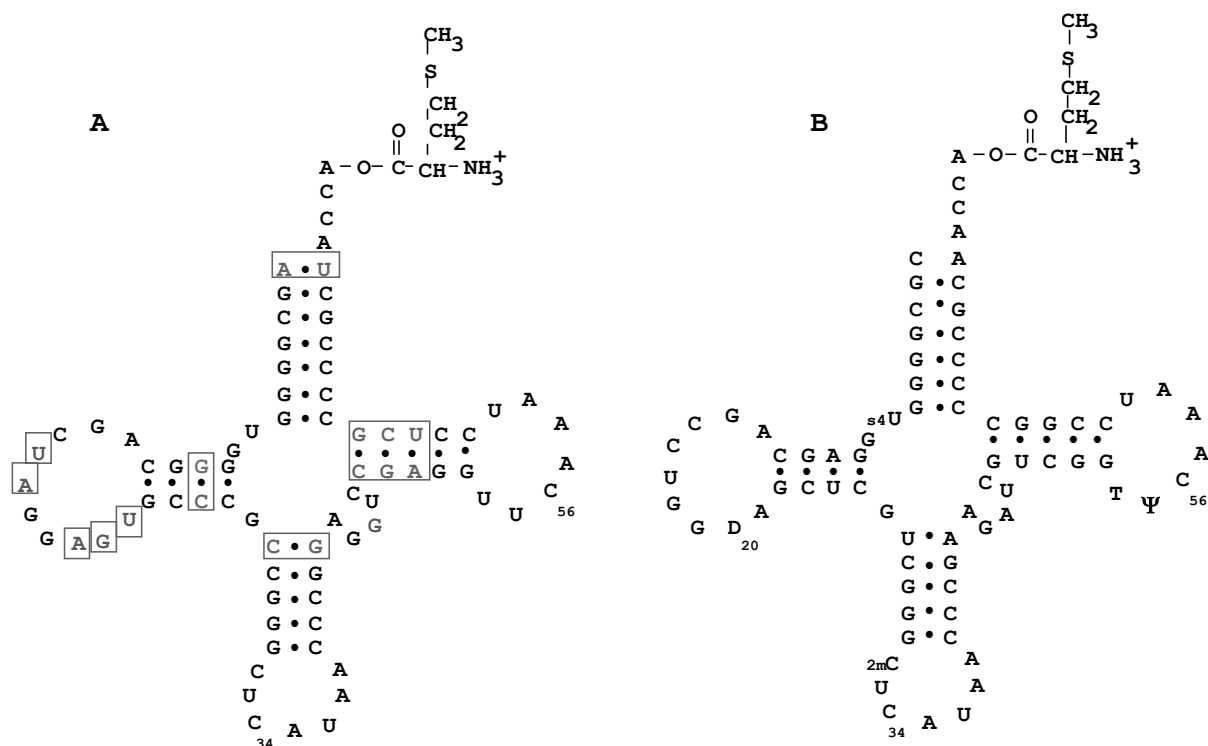
Préalablement, la vitesse de déacylation du [<sup>35</sup>S]-Met-ARNt<sub>i</sub><sup>Met</sup> a été mesurée en l'absence de Pa-aIF2. Ainsi, à 51°C, le Met-ARNt<sub>i</sub><sup>Met</sup> (80 nM) se déacyle à une vitesse de 0.28 min<sup>-1</sup>. En présence de 5 μM de Pa-aIF2 et de 1 mM de GDPNP, une forte protection contre la déacylation est observée : la vitesse de déacylation est en effet réduite à 0.009 min<sup>-1</sup>. En l'absence de GDPNP, cette protection est perdue et l'on retrouve une vitesse de 0.28 min<sup>-1</sup>. Par conséquent, un effet de protection similaire à celui d'EF1A est obtenu avec l'hétérotrimère Pa-aIF2 et il s'agit bien d'une interaction spécifique puisqu'en l'absence de GDPNP, et donc en l'absence de formation du complexe ternaire, l'effet de protection est annihilé. Lorsque le GDPNP est remplacé par du GDP (1mM), la vitesse de déacylation est réduite à 0.11 min<sup>-1</sup>. Ce résultat suggère qu'en présence de GDP, Pa-aIF2 est toujours susceptible de lier l'ARNt initiateur, bien que le complexe soit moins stable.

Ces résultats préliminaires nous ont donc permis de valider le test de protection mis en place et de l'utiliser pour mesurer les constantes de dissociation des complexes Pa-aIF2:ARNt.

Ainsi, un Kd de 90 ± 30 nM a été mesuré pour le complexe Pa-aIF2:Met-ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi* en présence de 1 mM GDPNP. Cette valeur est de plus inchangée si le GTP remplace son analogue non hydrolysable dans le test de protection. Ce résultat indique qu'aucune hydrolyse spontanée du GTP n'intervient au cours du test. La formation du complexe ternaire n'est donc pas suffisante pour entraîner l'hydrolyse du GTP. Celle-ci survient plus probablement dans le cadre de la sélection du bon codon de démarrage sur le ribosome, comme c'est le cas chez les Eucaryotes (Huang *et al.*, 1997). Une étude récente (Pedullà *et al.* 2005) a cependant montré que même en présence du ribosome et d'un oligonucléotide portant une séquence de type Shine-Dalgarno en amont d'un codon d'initiation AUG, l'hydrolyse du GTP n'a pas lieu. Ces résultats ont toutefois été obtenus en présence de complexes 70S, et par conséquent, dans un contexte où, a priori, l'hydrolyse du GTP porté par aIF2 a déjà eu lieu.

Pour déterminer si Pa-aIF2 interagit spécifiquement avec l'ARNt initiateur et si oui, quels sont les déterminants de l'ARNt reconnus par Pa-aIF2, il était crucial de disposer d'une panoplie d'ARNts mutés au niveau des positions potentielles de reconnaissance de l'ARNt par Pa-aIF2. Comme le souligne la **Figure 3.5**, l'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi* et l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* ont des séquences nucléotidiques très proches. En particulier, les tiges acceptrices de ces deux ARNts sont identiques à l'exception de la paire de base 1-72 : le faible appariement A<sub>1</sub>-U<sub>72</sub> présent chez l'ARNt de *P. abyssi* est remplacé par un mésappariement C<sub>1</sub>-A<sub>72</sub> chez l'ARNt bactérien. Il faut noter que les modifications post-traductionnelles de l'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi*

ne sont pas connues (Marck et Grosjean, 2002) et n'ont donc pas pu être représentées sur cette figure.

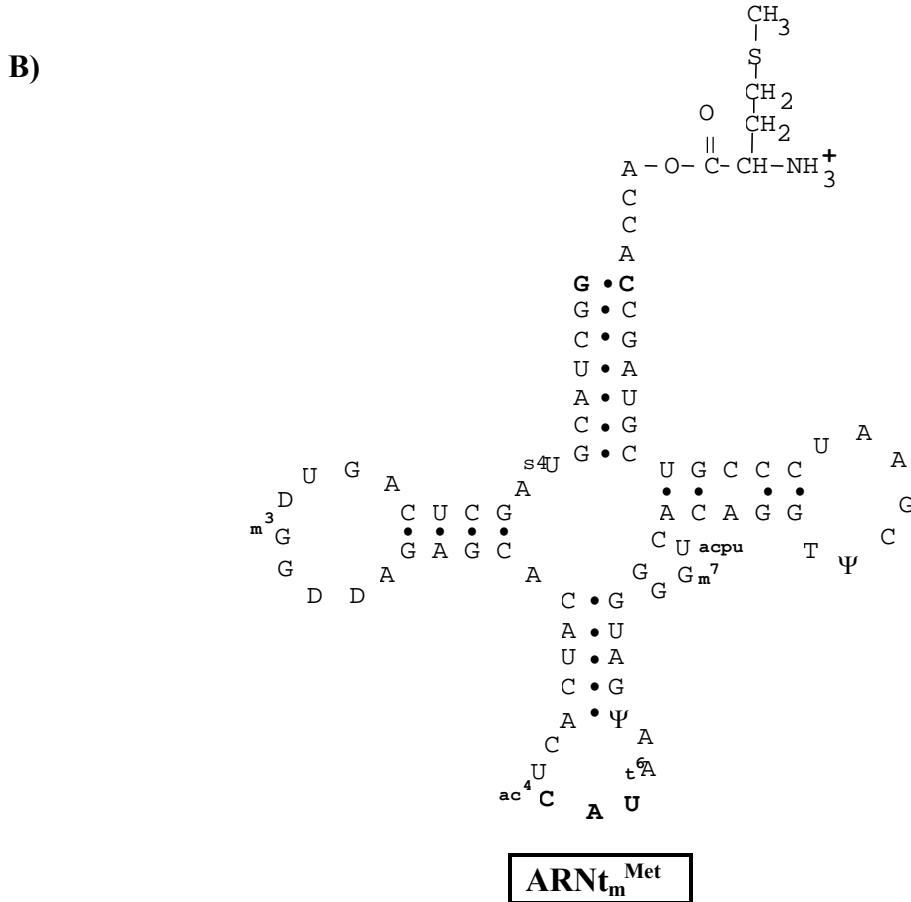
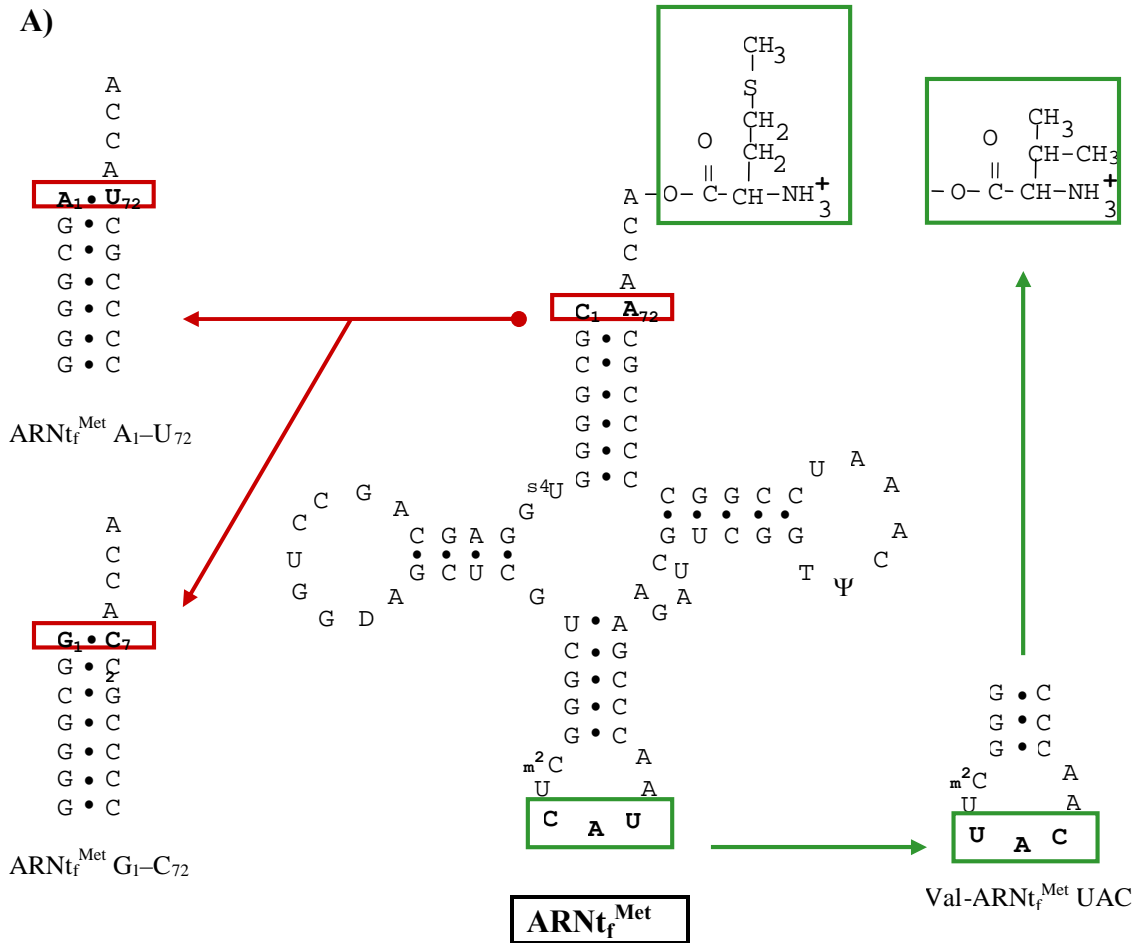


**Figure 3.5 :** Représentation en «feuille de trèfle» de l'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi* (A) et de l'ARNt<sub>i</sub><sup>Met</sup> d'*E. coli* (B). Les nucléotides qui diffèrent entre les deux ARNts sont encadrés et indiqués en gris sur la séquence de l'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi*.

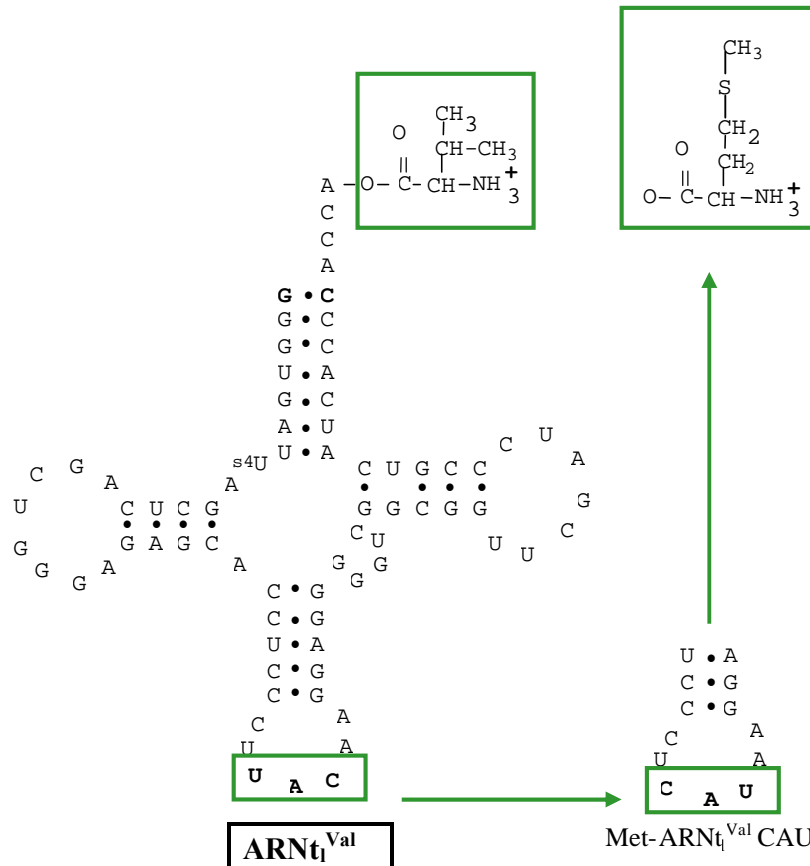
Cette forte similarité de séquence entre les deux ARNts nous a conduit à mesurer le Kd avec l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* aminoacylé avec de la [<sup>35</sup>S]-Met (10000 dpm/pmole). Un Kd de 120 ± 40 nM a été mesuré pour le complexe Pa-aIF2:Met-ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* en présence de 1 mM GDPNP. Ce résultat montre qu'en accord avec la ressemblance observée pour les deux ARNts, l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* est un aussi bon ligand de Pa-aIF2 que l'ARNt initiateur de *P. abyssi*. A partir de cette constatation, nous avons donc choisi d'utiliser l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* comme substrat modèle dans le test de protection. Ce choix nous a permis de tester les différents mutants de l'ARNt<sub>f</sub><sup>Met</sup> disponibles au laboratoire.

#### **4. Spécificité de Pa-aIF2 pour l'ARNt initiateur.**

Différents mutants de l'ARNt<sub>f</sub><sup>Met</sup> et de l'ARNt<sub>1</sub><sup>Val</sup> d'*E. coli*, construits et purifiés au laboratoire (Guillon *et al.*, 1992 ; Meinnel and Blanquet, 1995 ; Meinnel *et al.*, 1988), ainsi que l'ARNt<sub>m</sub><sup>Met</sup> élongateur d'*E. coli*, ont été testés pour leur capacité à être reconnus par le facteur d'initiation Pa-aIF2 de *P. abyssi*. Ces ARNts ont été purifiés et aminoacylés comme décrit dans la partie **Matériels et Méthodes**.



C)



**Figure 3.6 :** schéma des différents ARNts, natifs et mutants, utilisés dans le test de protection.

Ces mutants ont permis d'évaluer l'importance chez *P. abyssi* des principaux motifs caractéristiques de l'ARNt de démarrage de la traduction. Typiquement, deux zones ont été analysées : la paire de base 1-72 de la tige acceptrice, et l'acide aminé (avec mutation de l'anticodon). Les autres déterminants nucléotidiques ont été analysés de manière générale avec les différents ARNts à disposition. La **Figure 3.6** récapitule les différents ARNts et mutants utilisés.

Les différentes valeurs de K<sub>d</sub> mesurées avec ces ARNts sont indiquées dans le **Tableau 3.7**. Dans un premier temps, la spécificité de la réaction de protection de Pa-aIF2 vis-à-vis de l'ARNt initiateur a été testée en utilisant l'ARN<sub>f</sub><sup>Val</sup> élongateur valylé. Aucune protection n'est observée pour cet ARNt même en présence de 50 μM de Pa-aIF2, ce qui suggère bien une interaction spécifique de Pa-aIF2 avec l'ARNt initiateur. Les déterminants responsables de cette sélection spécifique ont ensuite été identifiés grâce à plusieurs mutants. Un ARN<sub>f</sub><sup>Met</sup> modifié par un anticodon UAC (qui code pour la valine) a d'abord été utilisé. Cet ARNt peut être aminoacylé avec une valine par la Valyl-ARNt synthétase (ValRS) d'*E. coli* (Guillon *et al.*, 1992 ; Schulman and Pelka, 1988). Le Val-ARN<sub>f</sub><sup>Met</sup> UAC résultant n'est

pas protégé de la déacylation spontanée par Pa-aIF2, même en présence de 50  $\mu\text{M}$  de Pa-aIF2. Ce résultat indique que le résidu méthionine est un déterminant essentiel de cette reconnaissance : le simple remplacement du résidu méthionine par une valine réduit l'affinité de l'hétérotrimère pour l'ARNt de plus de 2 ordres de grandeur.

	ARNt	Kd (nM)
1	Met-ARNt <sub>i</sub> <sup>Met</sup> <i>P. abyssii</i>	90 $\pm$ 30
2	Met-ARNt <sub>f</sub> <sup>Met</sup> <i>E. coli</i>	120 $\pm$ 40
3	Met-ARNt <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub> <i>E. coli</i>	90 $\pm$ 30
4	Met-ARNt <sub>m</sub> <sup>Met</sup> <i>E. coli</i>	450 $\pm$ 100
5	Met-ARNt <sub>i</sub> <sup>Val</sup> CAU <i>E. coli</i>	400 $\pm$ 100
6	Val-ARNt <sub>i</sub> <sup>Val</sup> <i>E. coli</i>	> 50 000
7	Val-ARNt <sub>f</sub> <sup>Met</sup> UAC <i>E. coli</i>	> 50 000
8	Met-ARNt <sub>f</sub> <sup>Met</sup> G <sub>1</sub> -C <sub>72</sub> <i>E. coli</i>	320 $\pm$ 80

**Tableau 3.7:** Constantes de dissociation des complexes réalisés entre Pa-aIF2 et les différents ARNts aminoacylés présents à une concentration de 80 nM.

Deux autres ARNts méthionylés ont été testés pour confirmer ce résultat : l'ARNt<sub>m</sub><sup>Met</sup> et un ARNt<sub>i</sub><sup>Val</sup> modifié par un anticodon CAU (qui code pour la méthionine). Ce dernier ARNt peut être aminoacylé avec une méthionine par la Methionyl-ARNt synthétase (MetRS) d'*E. coli* (Meinzel *et al.*, 1993). En présence de ces deux ARNts, des Kds de 450 nM et de 400 nM respectivement sont mesurés. Ces deux ARNts sont donc efficacement protégés par Pa-aIF2 bien que leur séquence nucléotidique diffère de celle de l'ARNt initiateur. Ces mesures confirment le rôle crucial joué par le résidu méthionine dans la reconnaissance de l'ARNt initiateur par Pa-aIF2. Cependant, les constantes de dissociation obtenues dans ce cas sont légèrement plus élevées (d'un facteur 3 à 4) que celle obtenue pour le Met-ARNt<sub>f</sub><sup>Met</sup>, ce qui suggère que la séquence polynucléotidique de l'ARNt participe également, bien que de façon minoritaire, à la discrimination positive de l'ARNt initiateur.

Dans ce contexte, nous nous sommes intéressés plus particulièrement à la paire de bases 1-72. Chez les Eucaryotes et chez les Archées, il s'agit d'une paire A<sub>1</sub>-U<sub>72</sub> qui est universellement conservée. Un rôle de cette paire de bases dans la reconnaissance spécifique de l'ARNt initiateur par eIF2 a déjà été postulé (Astrom *et al.*, 1993 ; Farruggio *et al.*, 1996). Pourtant, cette paire est remplacée par les bases non-appariées C<sub>1</sub> et A<sub>72</sub> dans l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* et celui-ci est efficacement protégé par Pa-aIF2. Le mutant Met-ARNt<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> présente un Kd du même ordre que celui de l'ARNt<sub>f</sub><sup>Met</sup> natif. Ainsi, la paire de bases A<sub>1</sub>-U<sub>72</sub> ne semble

pas être un déterminant essentiel malgré sa très haute conservation chez les Eucaryotes et chez les Archées. Par contre, le mutant Met-ARNt<sub>f</sub><sup>Met</sup> G<sub>1</sub>-C<sub>72</sub>, qui reproduit le fort appariement observé pour le Met-ARNt élongateur classique, présente un Kd de 320 nM, soit un Kd 3,5 plus élevé que celui de l'ARNt<sub>f</sub><sup>Met</sup> natif. L'absence d'un fort appariement au niveau de la paire de bases 1-72 contribue donc à une liaison optimale de l'ARNt par Pa-aIF2.

Finalement, nous avons constaté que ni l'ARNt<sub>f</sub><sup>Met</sup> non aminoacylé (jusqu'à une concentration de 10 µM) ni la méthionine seule (jusqu'à une concentration de 5 mM) n'inhibent la protection du Met-ARNt<sub>f</sub><sup>Met</sup> par Pa-aIF2.

Ainsi, l'essentiel de la reconnaissance de l'ARNt initiateur par le facteur aIF2 de *P. abyssi* porte plutôt sur la partie méthionine, la partie ribonucléique jouant cependant elle aussi un rôle discriminatoire, notamment au niveau de la paire de bases 1-72 pour laquelle un faible appariement est préféré.

#### **4) Rôle des différentes sous-unités de Pa-aIF2 dans la liaison de l'ARNt.**

##### **1. Rôle de Pa-aIF2γ.**

Après s'être intéressé au facteur Pa-aIF2 dans sa globalité, nous avons voulu déterminer le rôle précis de chacune des sous-unités de l'hétérotrimère dans cette liaison. Comme cela a déjà été souligné, Pa-aIF2γ présente une très forte homologie de structure avec le facteur d'élongation bactérien EF1A dans sa conformation active (EF1A:GTP). Cette ressemblance suggère un rôle central de la sous-unité γ dans la liaison à l'ARNt. La capacité de Pa-aIF2γ à lier l'ARNt initiateur a donc été testée à l'aide du test de protection contre la déacylation.

	ARNt	Kd (nM)
1	Met-ARNt <sub>f</sub> <sup>Met</sup> <i>E. coli</i>	5 000 ± 3 000
2	Val-ARNt <sub>f</sub> <sup>Met</sup> UAC <i>E. coli</i>	> 80 000
3	Met-ARNt <sub>i</sub> <sup>Val</sup> CAU <i>E. coli</i>	7 000 ± 3 000

**Tableau 3.8:** Constantes de dissociation des complexes réalisés entre Pa-aIF2γ et les différents ARNts aminoacylés présents à une concentration de 80 nM.

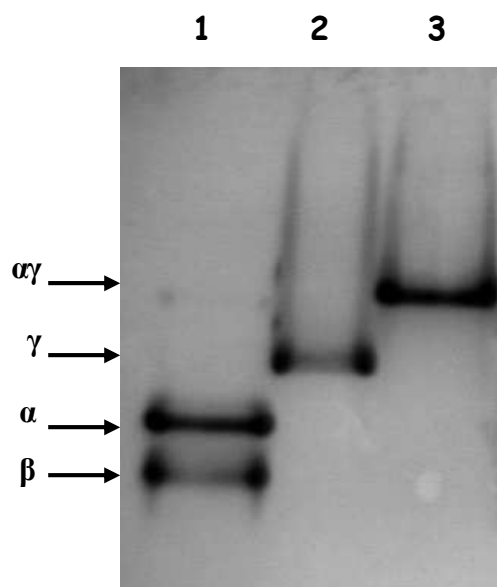
Effectivement, la sous-unité γ seule est capable de lier l'ARNt initiateur méthionylé avec une constante d'affinité de 5 µM (**Tableau 3.8**) en présence de 1 mM GDPNP. En

présence du Val-ARNt<sup>Met</sup> UAC, la capacité de protection de  $\gamma$  est perdue. Ainsi, Pa-aIF2 $\gamma$  seule est capable de reconnaître spécifiquement l'ARNt initiateur *via* le groupe méthionine. Ce résultat est confirmé puisqu'en présence du Met-ARNt<sup>Val</sup> CAU, la protection contre la déacylation est restaurée.

Cependant, l'affinité de la sous-unité  $\gamma$  seule pour l'ARNt initiateur est diminuée d'un facteur 50 par rapport à celle du facteur natif (Kd de 120 nM pour l'hétérotrimère). Cette large différence entre les deux constantes de dissociation montre qu'au moins une des deux autres sous-unités de l'hétérotrimère participe à la liaison de l'ARNt.

## 2. Rôle de Pa-aIF2 $\alpha$ .

Pour identifier la ou les autres sous-unités impliquées dans la liaison de Pa-aIF2 à l'ARNt, nous avons réitéré le test de protection avec les deux hétérodimères natifs. Comme pour l'hétérotrimère natif, les hétérodimères  $\alpha\gamma$  et  $\beta\gamma$  ont été obtenus en mélangeant dans un rapport équimolaire les sous-unités purifiées correspondantes et la stoechiométrie de l'assemblage a été vérifiée sur gel d'électrophorèse natif (**Figure 3.9**).



**Figure 3.9 :** Electrophorèse sur gel de polyacrylamide 12% natif acide pour tester la stoechiométrie du complexe Pa-aIF2 $\alpha\gamma$ . Piste 1 : sous-unités Pa-aIF2 $\alpha$  et Pa-aIF2 $\beta$  mélangées en quantités stoechiométriques → pas d'association ; Piste 2 : Pa-aIF2 $\gamma$  ; Piste 3 : sous-unités Pa-aIF2 $\alpha$  et Pa-aIF2 $\gamma$  mélangées en quantités stoechiométriques → une seule bande correspondant à l'hétérodimère est visible.

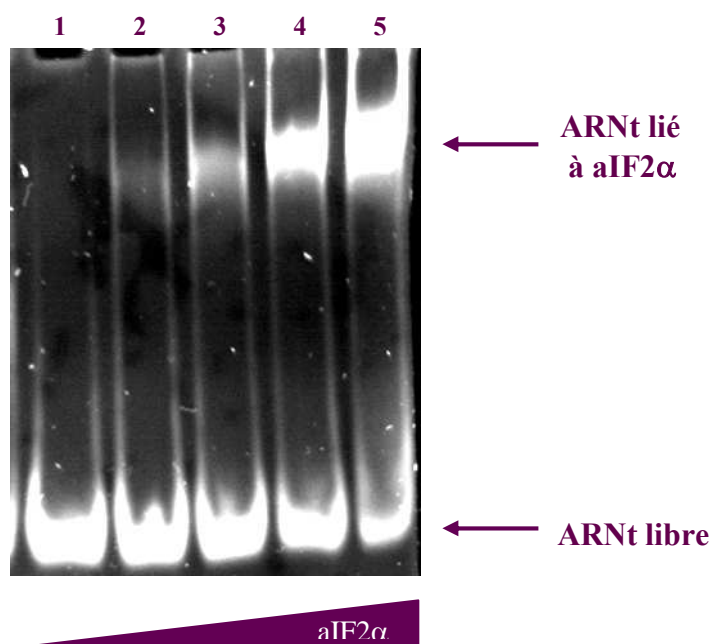
Les résultats pour les différents assemblages de Pa-aIF2 sont repris dans le **Tableau 3.10**. En présence de l'hétérodimère  $\beta\gamma$ , aucune augmentation de l'affinité pour l'ARNt<sup>Met</sup> n'est observée par rapport au monomère  $\gamma$ . Dans ce test de protection, la sous-unité  $\beta$  n'a donc pas de rôle dans la liaison de l'ARNt. Par contre, un Kd de 44 nM est obtenu avec

l'hétérodimère  $\alpha\gamma$  natif, soit une affinité comparable à celle de l'hétérotrimère natif. C'est donc Pa-aIF2 $\alpha$  qui participe, avec Pa-aIF2 $\gamma$ , à la liaison de l'ARNt.

	$\alpha$	$\beta$	$\gamma$	Kd (nM)
1	×	×	×	120 ± 40
2			×	5 000 ± 3 000
3		×	×	5 000 ± 3 000
4	×		×	44 ± 20
5	×			> 50 000
6		×		> 50 000

**Tableau 3.10 :** Constantes de dissociation des assemblages mentionnés en présence du Met-ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* 80 nM. Les sous-unités présentes dans l'assemblage testé sont indiquées par une croix.

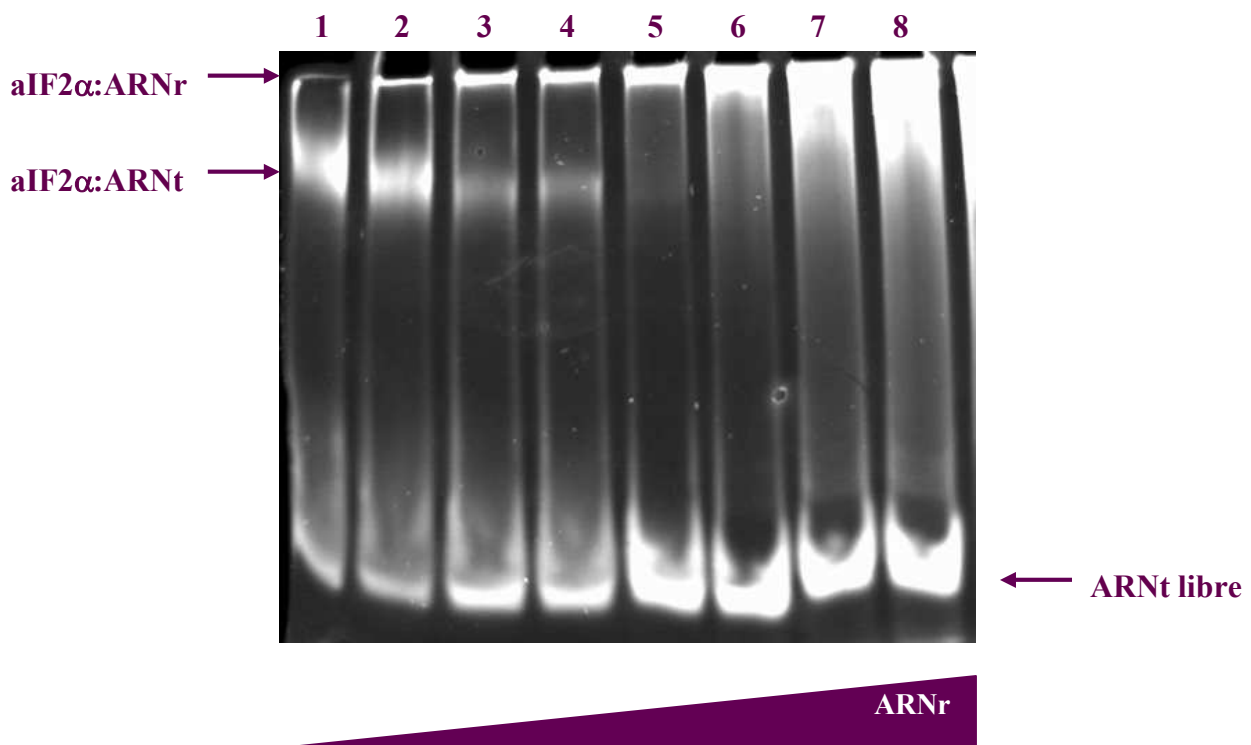
Des tests de protection contre la déacylation du Met-ARNt<sub>f</sub><sup>Met</sup> ont alors été réalisés en présence de 50  $\mu$ M du monomère  $\alpha$  : aucune protection n'est observée dans ce cas. Une interaction directe de Pa-aIF2 $\alpha$  avec l'ARNt n'est pas à exclure, mais elle n'impliquerait pas la partie 3' de l'ARNt.



**Figure 3.11 :** Expérience de retard de migration sur gel de l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* en présence d'aIF2 $\alpha$  de *P. abyssi*. 1.8  $\mu$ M d'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* ont été incubés dans 20  $\mu$ L de tampon 200 mM NaCl, 50 mM TBE pH 8.3, 10 mM 2-mercaptoéthanol, 1 mM MgCl<sub>2</sub>, 1 mM GDPNP en présence de concentrations croissantes de Pa-aIF2 $\alpha$  (0, 2.5, 5, 10 et 20  $\mu$ M de la Piste 1 à la Piste 5, respectivement). Après 15 minutes d'incubation à température ambiante, les échantillons sont chargés sur un gel à 5% en polyacrylamide contenant 50 mM TBE pH 8.3 et 1 mM MgCl<sub>2</sub>. La migration a lieu dans le même tampon. Le gel est ensuite révélé par coloration au BET. Un résultat identique est obtenu en présence d'ARNt<sub>f</sub><sup>Met</sup> aminoacylé.

Pour mieux comprendre le rôle joué par  $\alpha$ , des expériences de retard de migration sur gel de l'ARNt et/ou d'autres ARNs ont été réalisées en présence d' $\alpha$  (voir la partie **Matériels et Méthodes** pour le protocole détaillé). Pour cela, 1  $\mu\text{g}$  (1.8  $\mu\text{M}$ ) d'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* a été incubé en présence de concentrations croissantes de Pa-aIF2 $\alpha$  dans le tampon approprié. Les échantillons ont alors été déposés sur un gel d'électrophorèse natif. Après migration, le gel est plongé dans un bain de Bromure d'Ethidium pour révéler la migration de l'ARNt. La **Figure 3.11** illustre l'expérience.

La sous-unité  $\alpha$  seule est capable de retarder la migration de l'ARNt sur gel. Dans les mêmes conditions, aucun retard de migration de l'ARNt n'a été observé avec la sous-unité  $\beta$  et/ou  $\gamma$ . Cependant, la capacité de liaison des ARNs de Pa-aIF2 $\alpha$  n'est pas spécifique de l'ARNt. En effet, l'expérience a été reproduite en présence de quantités croissantes d'un mélange d'ARN ribosomaux 16S et 23S d'*E. coli*. Comme le montre la **Figure 3.12**, ce mélange d'ARNr entre en compétition avec l'ARNt pour la liaison à  $\alpha$  et en présence d'un rapport massique 2:1 de l'ARN ribosomal par rapport à l'ARNt, la liaison de l'ARNt par  $\alpha$  est abolie (Piste 5).



**Figure 3.12 :** Compétition du mélange ARNr 16S/23S et de l'ARNt pour la liaison à Pa-aIF2 $\alpha$ . 13.6  $\mu\text{g}$  (20  $\mu\text{M}$ ) de Pa-aIF2 $\alpha$  de *P. abyssi* et 1  $\mu\text{g}$  (1.8  $\mu\text{M}$ ) d'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* ont été incubés dans 20  $\mu\text{L}$  de tampon 200 mM NaCl, 50 mM TBE pH 8.3, 10 mM 2-mercaptoéthanol, 1 mM MgCl<sub>2</sub>, 1 mM GDPNP. Des quantités croissantes du mélange d'ARNr 16S et 23S d'*E. coli* (0, 0.2, 0.5, 1, 2, 4, 7.5 et 10  $\mu\text{g}$  de la Piste 1 à la Piste 8, respectivement) ont été ajoutées dans le mélange. Les échantillons ont ensuite été traités comme pour la **Figure 3.11**.

Ainsi, la sous-unité  $\alpha$  seule a la capacité de se lier aux acides nucléiques. Cette propriété pourrait fournir une explication au rôle joué par Pa-aIF2 $\alpha$  dans la liaison de l'ARNt initiateur par l'hétérotrimère. Cependant, cette affinité n'étant pas spécifique du repliement de type ARNt, aucune conclusion quant à une interaction directe de Pa-aIF2 $\alpha$  avec l'ARNt ne peut être déduite.

## **Partie B : Cartographie des zones d'interactions entre les sous-unités du facteur Pa-aIF2.**

### **La sous-unité $\gamma$ , maillon central de l'assemblage.**

L'étude de l'affinité du facteur Pa-aIF2 pour l'ARNt initiateur a mis en exergue une nouvelle entité au sein de l'hétérotrimère : l'hétérodimère  $\alpha\gamma$ . En effet, la sous-unité  $\beta$  ne semble pas jouer de rôle dans la liaison au principal ligand de Pa-aIF2, le Met-ARNt<sub>i</sub><sup>Met</sup>. Une nouvelle géographie se dessine donc autour de l'ARNt : la sous-unité  $\gamma$  interagit directement avec l'ARNt, et la sous-unité  $\alpha$ , qui apporte un gain d'affinité considérable, doit donc être assez proche du site de fixation de la tige acceptrice de l'ARNt sur  $\gamma$ . Au contraire, on imagine une position plus éloignée pour la sous-unité  $\beta$ . Ces hypothèses nous ont donc tout naturellement conduit vers l'étude des interactions entre les différentes sous-unités de Pa-aIF2, pour tenter de comprendre l'agencement structural qui favorise la formation du complexe ternaire.

#### **1) Obtention des mutants de la sous-unité $\gamma$ .**

##### **1. Obtention des différents plasmides mutés.**

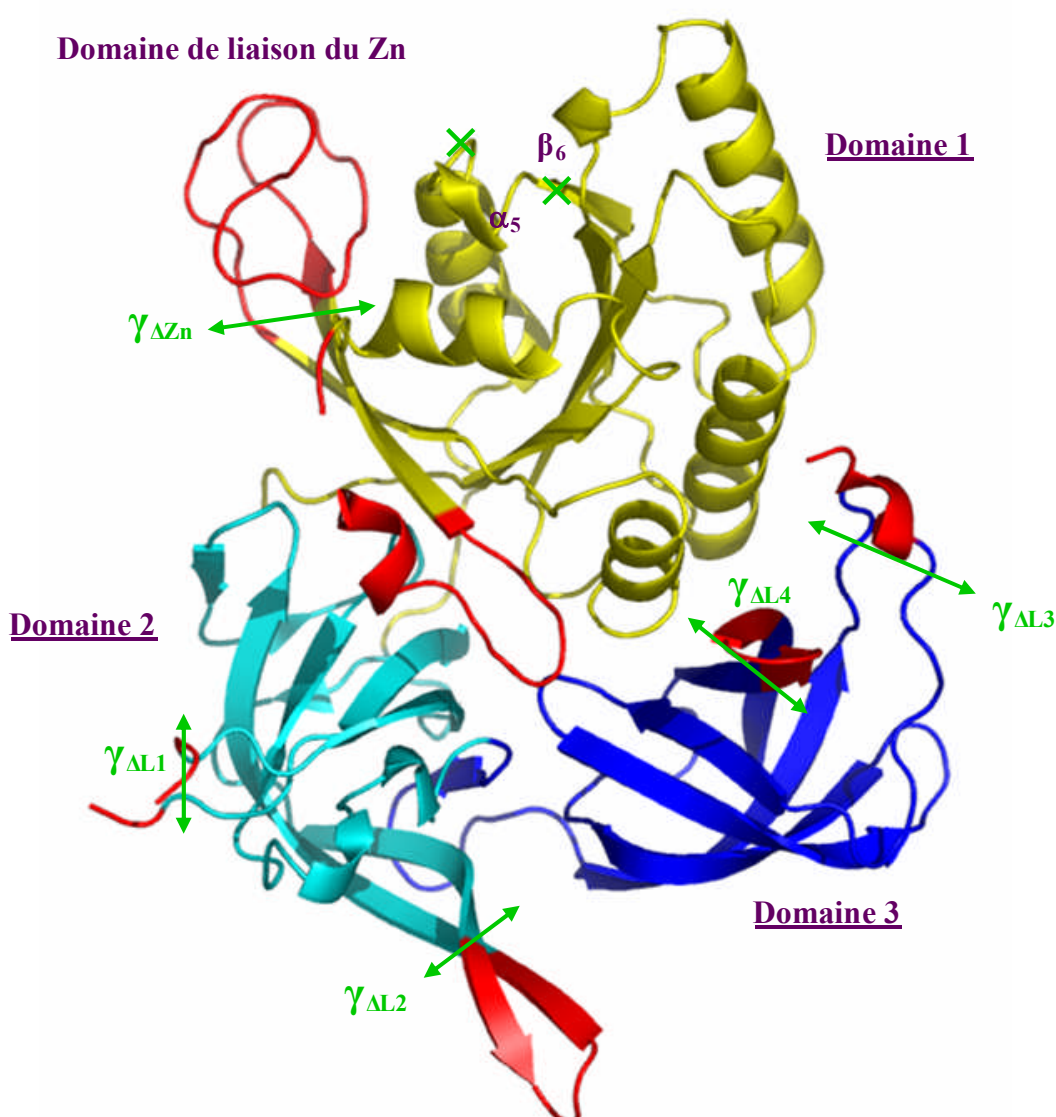
Les zones caractéristiques de Pa-aIF2 $\gamma$  identifiées par comparaison avec EF1A sont représentées sur la **Figure 3.13**. Elles ont fait l'objet d'une étude fonctionnelle par mutagenèse dirigée. En plus des boucles caractéristiques d'aIF2 $\gamma$  décrites au **Chapitre II**, deux résidus hautement conservés dans les séquences d'e/aIF2 $\gamma$  ont été envisagés. Ces résidus P178 et N186 se trouvent dans une zone au sommet du domaine 1 correspondant au brin  $\beta 6$  et à la petite hélice  $\alpha 5$  (voir **Annexe 1**). Chez EF1A, ils ne sont pas conservés et la zone correspond à une zone d'insertion par rapport à Pa-aIF2 $\gamma$  dans la séquence de certains EF1A. De plus, ces résidus sont proches du domaine de fixation du zinc, qui pourrait avoir lui-même un rôle fonctionnel non négligeable. Enfin, ils se trouvent à la surface de la protéine, tournés vers le solvant, et donc potentiellement accessibles par un éventuel partenaire d'interaction.

Pour les boucles L1, L2, L3, L4 ainsi que pour le domaine du zinc, des mutants de délétions ont été envisagés (Voir **Figure 3.13**). Pour la zone englobant le brin  $\beta 6$  et l'hélice  $\alpha 5$ , moins caractéristique que les régions précédemment citées, seuls deux mutants ponctuels ont été envisagés :  $\gamma_{P178A}$  et  $\gamma_{N186A}$ . Enfin, des mutants ponctuels additionnels ont été utilisés pour étudier le rôle du domaine de fixation du zinc :  $\gamma_{C60A}$ ,  $\gamma_{C63H}$  et  $\gamma_{C72A}$ .

Les mutants  $\gamma_{\Delta L1}$ ,  $\gamma_{\Delta L2}$ ,  $\gamma_{\Delta Zn}$ ,  $\gamma_{C60A}$  et  $\gamma_{C72A}$ , ont été obtenus par mutagenèse dirigée par PCR dite de « chevauchement » (voir **Matériels et Méthodes** pour le principe de la PCR).

Les mutants  $\gamma_{\Delta L3}$ ,  $\gamma_{\Delta L4}$ ,  $\gamma_{C63H}$ ,  $\gamma_{P178A}$  et  $\gamma_{N186A}$ , ont été obtenus par mutagenèse dirigée par PCR en une seule étape à l'aide de la technique *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene (voir **Matériel et Méthodes**).

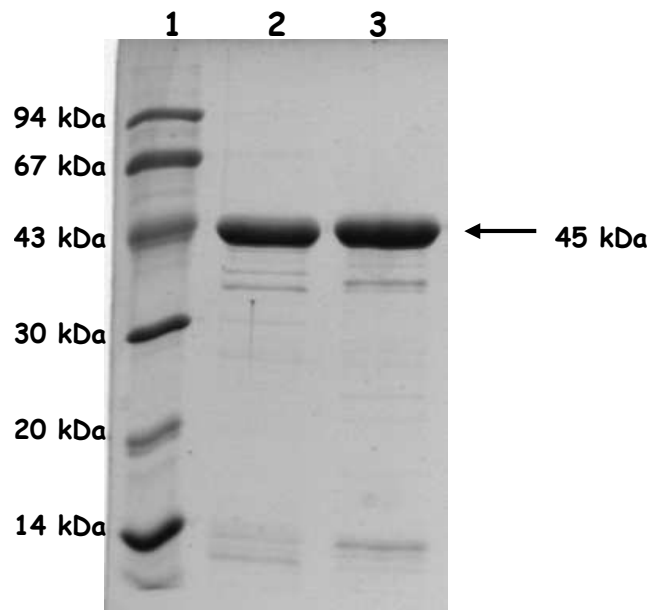
Les dix plasmides mutants obtenus ont été séquencés dans chaque cas pour vérifier la conformité de la construction.



**Figure 3.13 :** Structure de Pa-aIF2 $\gamma$  (Schmitt *et al.*, 2002). Les principales zones de divergence avec EF-Tu:GDPNP ont été représentées en rouge. Les flèches vertes désignent les différents mutants de délétion de la sous-unité  $\gamma$  qui ont été construits au cours de cette étude.

## 2. Production et purification des mutants de Pa-aIF2 $\gamma$ .

Les plasmides mutants ont été utilisés pour transformer des cellules BL21(DE3) d'*E. coli* en présence du plasmide pSJS1240 et les protéines mutées ont été surproduites de la même manière que la protéine native. A l'exception des mutants  $\gamma_{\Delta Zn}$ ,  $\gamma_{C60A}$  et  $\gamma_{C72A}$ , toutes ces protéines ont été purifiées selon le même protocole que celui utilisé pour la sous-unité  $\gamma$  native (**Figure 3.14**). De bons rendements de production ont été obtenus avec ces mutants, souvent même supérieurs à ceux obtenus pour la protéine native.



**Figure 3.14 :** Purification des mutants de délétion de la sous-unité  $\gamma$  de *P. abyssi*. Analyse des échantillons purifiés par migration électrophorétique sur gel de polyacrylamide (12%) dénaturant. Piste 1 : marqueur de poids moléculaire ; Piste 2 :  $\gamma_{\Delta L1}$ ; Piste 3 :  $\gamma_{\Delta L2}$ .

## 3. Production et purification des mutants du domaine du zinc.

Pour les trois mutants  $\gamma_{\Delta Zn}$ ,  $\gamma_{C60A}$  et  $\gamma_{C72A}$ , un très faible taux de surproduction a été obtenu. Ceci est très probablement dû au fait qu'ils affectent la coordination du zinc à  $\gamma$ . Ces résultats suggèrent un rôle important du domaine de liaison du zinc dans la stabilité de Pa-aIF2 $\gamma$ . L'analyse des taux de surproduction de ces différents mutants sur gel d'électrophorèse en conditions dénaturantes a d'ailleurs révélé, dans certains cas, la présence d'une bande autour de 45 kDa qui pourrait correspondre aux Pa-aIF2 $\gamma$  mutants dans la fraction insoluble obtenue après sonication des cultures bactériennes. La présence de ces Pa-aIF2 $\gamma$  mutés dans la fraction insoluble témoignerait ainsi d'un mauvais repliement des protéines dû à l'altération de leur domaine de fixation du zinc. Enfin, lorsqu'un mutant a finalement pu être légèrement surproduit et obtenu dans la fraction soluble au cours de ces diverses tentatives, une

thermosensibilité de la protéine à 80°C a été observée. Dans ce cas, la purification de la protéine correspondante a été entreprise en éliminant l'étape de chauffage à 80°C. Cependant, la protéine d'intérêt étant trop peu surproduite et les protéines contaminantes d'*E. coli* n'ayant pas été éliminées par chauffage, la méthode de purification classique de Pa-aIF2 $\gamma$  n'a finalement pas permis d'obtenir une protéine  $\gamma$  mutée purifiée.

A contrario, le mutant  $\gamma_{C63H}$  ne présente pas ce défaut de surproduction. On peut supposer que la mutation conservatrice de la cystéine en une histidine, résidu également bon chélateur de cations, n'entraîne pas la perte d'un degré de coordination autour de l'atome de zinc.

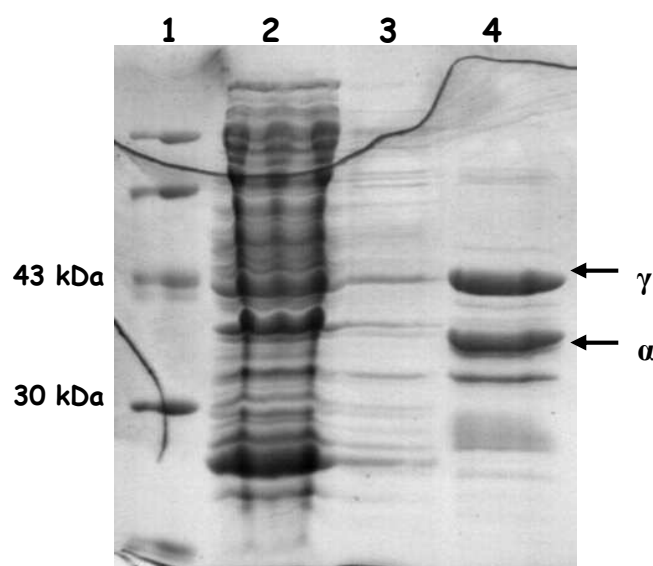
Pour essayer de pallier ce défaut de surproduction, différentes tentatives ont été entreprises :

- les conditions de surproduction ont été modifiées en ajoutant du zinc dans le milieu de croissance des bactéries ou bien en diminuant la température et/ou la quantité d'IPTG lors de la phase d'induction, ceci dans le but de diminuer la fraction de protéine insoluble qui pourrait se retrouver piégée dans des corps d'inclusion.
- une autre souche bactérienne, plus performante, a été utilisée : la souche Rosetta (DE3) pLacI RARE d'*Escherichia coli*. Cette souche contient un plasmide qui permet d'exprimer les ARNts rares ArgU, IleY et LeuW, capables de déchiffrer les codons rares rencontrés sur les ARNms archéens, ainsi que lacI. Son utilisation évite ainsi la co-transformation du pET3alpa portant l'insert muté avec un vecteur de type pSJS1240 ou pSBETa. De plus, cette souche est également mutée dans le gène de la perméase au lactose, ce qui rend l'induction à l'IPTG plus contrôlable.
- enfin, lors des tests de purification sur les échantillons légèrement surproduits, du zinc a été ajouté dans toutes les étapes pour tenter de renforcer la stabilité de la protéine.

Malgré ces différents essais, la technique de purification employée pour la sous-unité  $\gamma$  native (avec quatre colonnes chromatographiques successives) reste peu sélective pour ces mutants faiblement surproduits. Une nouvelle méthode de purification a donc été mise au point. Une technique efficace aurait pu consister à sélectionner une colonne d'affinité qui retiendrait spécifiquement la protéine  $\gamma$ . Malheureusement, la protéine Pa-aIF2 $\gamma$  étiquetée par 6 His en N-terminal n'est pas retenue sur une colonne d'affinité de type TALON (Clontech). Par contre, l'hétérodimère  $\alpha\gamma$  où  $\alpha$  porte une queue polyhistidine en N-terminal est retenu sur une telle colonne d'affinité. Ce choix de purification a été validé par les essais préliminaires d'assemblage avec la sous-unité  $\gamma$  réalisés sur les extraits protéiques bruts obtenus après chauffage à 80°C des lysats cellulaires. Ces premiers essais semblent indiquer que les mutants

du domaine de fixation du zinc n'interviennent pas dans la liaison à  $\alpha$ . Le domaine du zinc pourrait cependant jouer un rôle au niveau de la liaison à  $\beta$ . Ces premiers tests ont donc permis de légitimer le choix de purification de l'hétérodimère  $\alpha\gamma$ .

Le principe général de la purification est repris dans la partie **Matériels et Méthodes**. Les trois mutants du domaine de fixation du zinc ont donc été purifiés sous forme d'hétérodimères  $\alpha\gamma$ . Dans un premier temps, la sous-unité  $\alpha$  isolée a été préparée. Puis, un aliquot de la protéine purifiée a été ajouté dans l'extrait protéique brut de Pa-aIF2 $\gamma$  mutant obtenu après lyse cellulaire et chauffage à 80°C. La quantité de Pa-aIF2 $\alpha$  ajoutée est telle que la protéine  $\alpha$  reste en défaut par rapport à  $\gamma$ , de manière à ce que seuls les hétérodimères parfaitement stœchiométriques soient retenus sur la colonne TALON. Les hétérodimères  $\alpha\gamma_{\text{mutant}}$  ainsi purifiés présentent une bonne pureté et la quantité obtenue permet d'envisager une étude biochimique. La **Figure 3.15** présente la purification de l'hétérodimère  $\alpha\gamma_{\Delta\text{Zn}}$  sur colonne d'affinité TALON à base d'ions  $\text{Co}^{2+}$ .

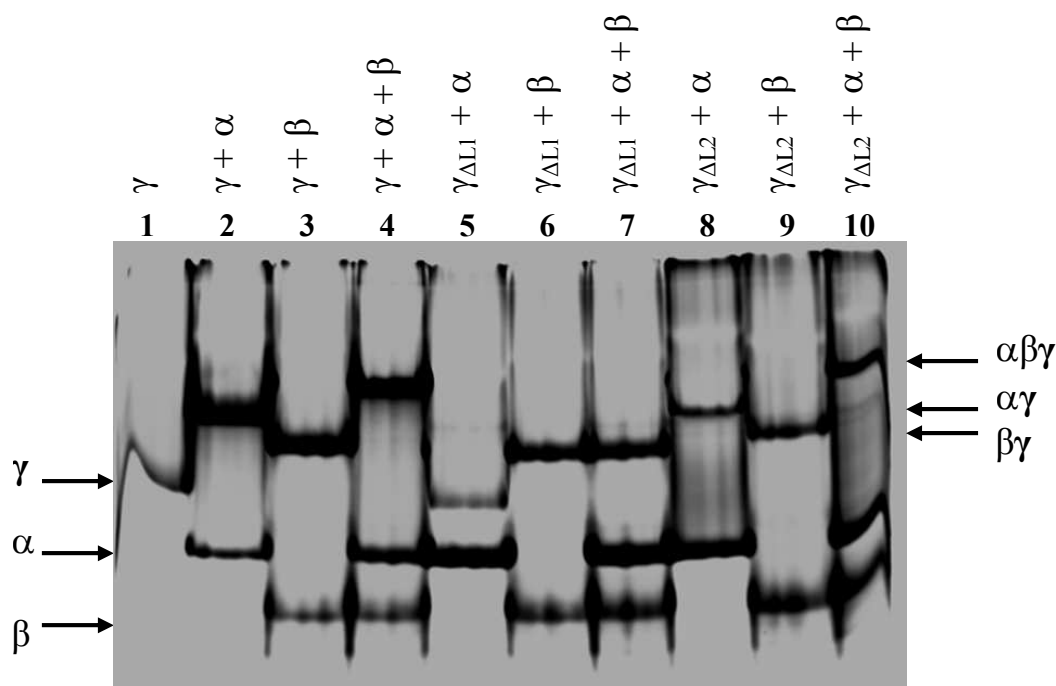


**Figure 3.15 :** Purification de l'hétérodimère  $\alpha\gamma_{\Delta\text{Zn}}$  de *P. abyssi* sur colonne d'affinité TALON. Analyse des échantillons purifiés par migration électrophorétique sur gel de polyacrylamide 12 % dénaturant. Piste 1 : marqueur de poids moléculaire ; Piste 2 : aliquot des fractions de tête de la TALON ; Piste 3 : aliquot des fractions de lavage à 10 mM imidazole de la TALON ; Piste 4 : aliquot des fractions d'élution de la TALON, contenant l'hétérodimère purifié.

## **2) Tests d'assemblages sur gels natifs.**

Après avoir obtenu les différents mutants de la sous-unité  $\gamma$  de *P. abyssi*, nous avons testé leur capacité à lier la sous-unité  $\alpha$  et/ou la sous-unité  $\beta$  sur gels d'électrophorèse en conditions non dénaturantes.

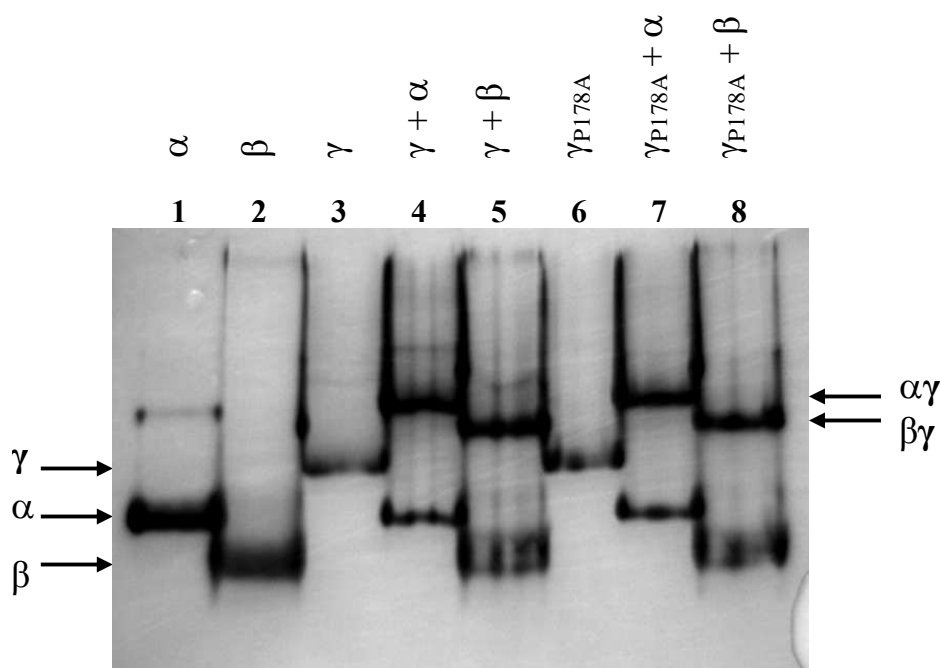
Du fait des points isoélectriques basiques des protéines aIF2 de *P. abyssi*, des gels de polyacrylamide acides ont été préparés. Il s'agit de gels à 12% en polyacrylamide dans un tampon contenant un mélange 240 mM KOH/144 mM d'acide acétique de pH 4.3. Les assemblages hétérodimériques et hétérotrimériques des sous-unités  $\alpha$  et  $\beta$  avec les différents mutants  $\gamma$  ont été réalisés dans un volume final de 20  $\mu$ L, en introduisant toujours la sous-unité  $\gamma$  en défaut par rapport aux autres sous-unités. Ceci permet de visualiser si toute la protéine  $\gamma$  introduite dans le mélange est « consommée » pour former l'assemblage étudié ou si, du fait d'un défaut d'assemblage, une fraction de  $\gamma$  reste libre. Les **Figures 3.16** et **3.17** illustrent les gels respectivement obtenus en présence de mutants de boucles (**Figure 3.16**) et d'un mutant ponctuel (**Figure 3.17**). Le même type de gel a été reproduit avec les différents mutants.



**Figure 3.16 :** Assemblage des sous-unités  $\alpha$  et  $\beta$  natives de Pa-aIF2 avec les mutants de boucles  $\gamma_{\Delta L1}$  et  $\gamma_{\Delta L2}$  de Pa-aIF2 $\gamma$ . Les protéines indiquées au-dessus de chaque piste ont été préalablement mélangées dans 10  $\mu$ L de tampon 200 mM NaCl, 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol. 10  $\mu$ L de tampon de dépôt (240 mM KOH, 144 mM acide acétique, 20% glycérol, 0.4% (w/v) bleu de méthylène) ont été ajoutés et les échantillons ont été déposés sur un gel natif acide (12% en polyacrylamide). Après migration, les gels ont été colorés au bleu de Coomassie. Pistes 1 à 4 :  $\gamma$  (12  $\mu$ M, piste 1) mélangé avec un excès molaire d' $\alpha$  (piste 2), de  $\beta$  (piste 3) ou d' $\alpha+\beta$  (piste 4) ; Pistes 5 à 7 :  $\gamma_{\Delta L1}$  (15  $\mu$ M) mélangé avec un excès molaire d' $\alpha$  (piste 5), de  $\beta$  (piste 6) ou d' $\alpha+\beta$  (piste 7) ; Pistes 8 à 10 :  $\gamma_{\Delta L2}$  (15  $\mu$ M) mélangé avec un excès molaire d' $\alpha$  (piste 8), de  $\beta$  (piste 9) ou d' $\alpha+\beta$  (piste 10).

Sur la **Figure 3.16**, les bandes relatives aux deux hétérodimères  $\alpha\gamma_{\Delta L2}$  (piste 8) et  $\beta\gamma_{\Delta L2}$  (piste 9) apparaissent très clairement par comparaison avec les hétérodimères natifs

(pistes 2 et 3). De même, l'hétérotrimère  $\alpha\beta\gamma_{\Delta L2}$  est formé (piste 10). Par contre, les deux sous-unités  $\alpha$  et  $\gamma_{\Delta L1}$  restent dissociées (piste 5). L'hétérodimère  $\beta\gamma_{\Delta L1}$  est formé comme en témoigne la bande supérieure de la piste 6. Enfin, dans le cas de l'hétérotrimère (piste 7), la sous-unité  $\alpha$  reste encore une fois dissociée : seul l'hétérodimère  $\beta\gamma_{\Delta L1}$  est formé, la sous-unité  $\alpha$  n'est plus liée par le mutant  $\gamma$ . Cette absence de liaison de la sous-unité  $\alpha$  par le mutant  $\gamma_{\Delta L1}$  suggère que la suppression de la boucle L1 annihile la liaison de  $\gamma$  à  $\alpha$ . La boucle L1 du domaine 2 de Pa-aIF2 $\gamma$  est par conséquent clairement impliquée dans la fixation de la sous-unité Pa-aIF2 $\alpha$ .



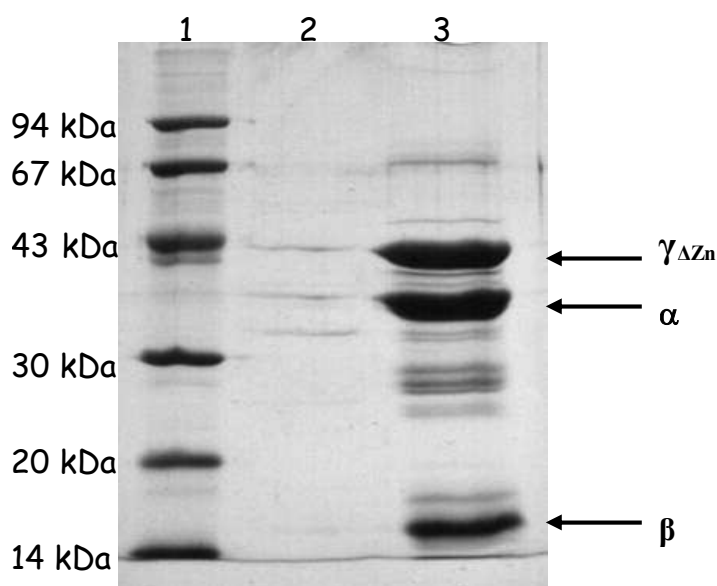
**Figure 3.17:** Assemblage des sous-unités  $\alpha$  et  $\beta$  natives de Pa-aIF2 avec le mutant ponctuel  $\gamma_{P178A}$  de Pa-aIF2 $\gamma$ . Les protéines indiquées au-dessus de chaque piste ont été préalablement mélangées dans 10  $\mu$ L de tampon 200 mM NaCl, 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol. 10  $\mu$ L de tampon de dépôt (240 mM KOH, 144 mM acide acétique, 20% glycérol, 0.4% (w/v) bleu de méthylène) ont été ajoutés et les échantillons ont été déposés sur un gel natif acide (12% en polyacrylamide). Après migration, les gels ont été colorés au bleu de Coomassie. Piste 1 :  $\alpha$  (24  $\mu$ M); Piste 2 :  $\beta$  (24  $\mu$ M); Pistes 3 à 5 :  $\gamma$  (12  $\mu$ M, piste 3) mélangé avec un excès molaire d' $\alpha$  (piste 4) ou de  $\beta$  (piste 5); Pistes 6 à 8 :  $\gamma_{P178A}$  (12  $\mu$ M, piste 6) mélangé avec un excès molaire d' $\alpha$  (piste 7) ou de  $\beta$  (piste 8).

La **Figure 3.17** illustre la formation des deux hétérodimères  $\alpha\gamma_{P178A}$  et  $\beta\gamma_{P178A}$ . Le résidu P178 ne semble donc pas avoir de rôle particulier dans la liaison à l'une ou l'autre des deux sous-unités  $\alpha$  et  $\beta$  de Pa-aIF2.

Pour les mutants relatifs au domaine du zinc, et en particulier pour le mutant  $\gamma_{\Delta Znp}$ , la migration sur gels natifs de polyacrylamide est moins nette. Les résultats obtenus par cette

méthode semblent toutefois indiquer que l'association à  $\beta$  n'est pas amoindrie en l'absence du domaine du zinc. Cependant, une confirmation de ce résultat était nécessaire. Pour cela, nous avons regardé l'assemblage de  $\beta$  avec l'hétérodimère purifié  $\alpha\gamma_{\Delta Zn}$  directement sur la colonne d'affinité TALON.

Ainsi, la protéine  $\beta$  purifiée de manière classique (selon le protocole décrit dans la Partie A de ce chapitre) a été ajoutée à un aliquot du complexe purifié  $\alpha\gamma_{\Delta Zn}$ . Le mélange a été réalisé de manière à ce que la protéine  $\beta$  soit en léger défaut molaire par rapport aux deux autres sous-unités. Ainsi, si toute la protéine se fixe à la sous-unité  $\gamma$ , sans défaut d'association dû à la mutation, alors il ne devrait plus rester de  $\beta$  dans les fractions de tête de la colonne : toute la protéine reste accrochée sur la résine sous forme  $\alpha\beta\gamma_{\Delta Zn}$ . Une fois le mélange injecté sur la colonne, la purification a été poursuivie comme pour l'hétérodimère : après passage des fractions de tête, la colonne a été rincée avec du tampon contenant 10 mM imidazole puis le complexe est élué avec un tampon contenant 250 mM imidazole. Le contenu des pics correspondant aux différentes phases de purification a été analysé sur gel SDS-PAGE. La **Figure 3.18** présente le gel obtenu.



**Figure 3.18 :** Analyse de la purification du complexe  $\alpha\gamma_{\Delta Zn}$  en présence de la sous-unité  $\beta$ . Les fractions des pics correspondant à chaque étape de purification (injection, rinçage, élution) ont été rassemblées et un aliquot de 20  $\mu$ L de chaque échantillon a été déposé sur un gel à 12% en polyacrylamide. Après migration, le gel a été coloré au Bleu de Coomassie. Piste 1 : marqueur de poids moléculaire ; Piste 2 : aliquot du pic d'injection ; Piste 3 : aliquot du pic d'élution. Comme le test d'assemblage a été réalisé avec des échantillons déjà purifiés, les contaminants sont minimes et aucun pic n'a été obtenu lors de la phase de rinçage à 10 mM imidazole. L'aliquot correspondant n'est donc pas indiqué.

Ce gel montre clairement que toute la protéine  $\beta$  initialement introduite s'accroche à l'hétérodimère  $\alpha\gamma_{\Delta Zn}$  par l'intermédiaire de la sous-unité  $\gamma$  puisqu' $\alpha$  et  $\beta$  n'interagissent pas ensemble. Par conséquent, la suppression du domaine de fixation du zinc ne semble pas affecter la liaison de la sous-unité  $\beta$  à  $\gamma$ .

Cette méthode d'assemblage sur gel natif nous a donc permis d'étudier le rôle des différentes positions caractéristiques de Pa-aIF2 $\gamma$  identifiées par comparaison avec EF1A. Le **Tableau 3.19** récapitule les résultats obtenus pour les différents mutants testés.

Mutants $\gamma$	Liaison à $\alpha$	Liaison à $\beta$	Formation de l'hétérotrimère
$\gamma_{\Delta L1}$		✗	Non, seul $\beta\gamma$ est formé
$\gamma_{\Delta L2}$	✗	✗	Oui
$\gamma_{\Delta L3}$	✗	✗	Oui
$\gamma_{\Delta L4}$	✗	✗	Oui
$\gamma_{P178A}$	✗	✗	Oui
$\gamma_{N186A}$	✗	✗	Oui
$\gamma_{C63H}$	✗	✗	Oui
$\gamma_{\Delta Zn}$	✗	✗	Oui
$\gamma_{\Delta C72A}$	✗	✗	Oui
$\gamma_{\Delta C60A}$	✗	✗	Oui

**Tableau 3.19 :** Récapitulatif des test d'assemblages des sous-unités  $\alpha$  et  $\beta$  réalisés en présence des différents mutants de  $\gamma$  par électrophorèse sur gel natif. Une croix indique lorsque l'assemblage entre les protéines étudiées est réalisé. Enfin, la dernière colonne reprend les résultats observés pour la formation du hétérotrimère.

Toutes ces expériences d'assemblage par électrophorèse native ont été confirmées par des tests d'assemblages réalisés par chromatographie d'exclusion. Le protocole est détaillé dans la partie **Matériels et Méthodes**. Brièvement, les différentes sous-unités purifiées, natives et mutées, ont été mélangées dans un tampon 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol, 500 mM NaCl. Les protéines individuelles ainsi que les différents assemblages à tester ont ensuite été chromatographiés sur une colonne d'exclusion de type TSK3000 qui permet de séparer les composés en fonction de leur taille. Le passage des monomères et des assemblages natifs a permis de calibrer la colonne et de connaître ainsi les temps de rétention caractéristiques de chaque forme de Pa-aIF2. Ce calibrage a alors permis de déterminer pour chaque assemblage si les protéines en présence s'associaient ou non, en

fonction des temps de rétention obtenus. La **Figure 3.27** illustre le type de chromatogrammes obtenus avec l'assemblage  $\alpha_{D3}\gamma$ .

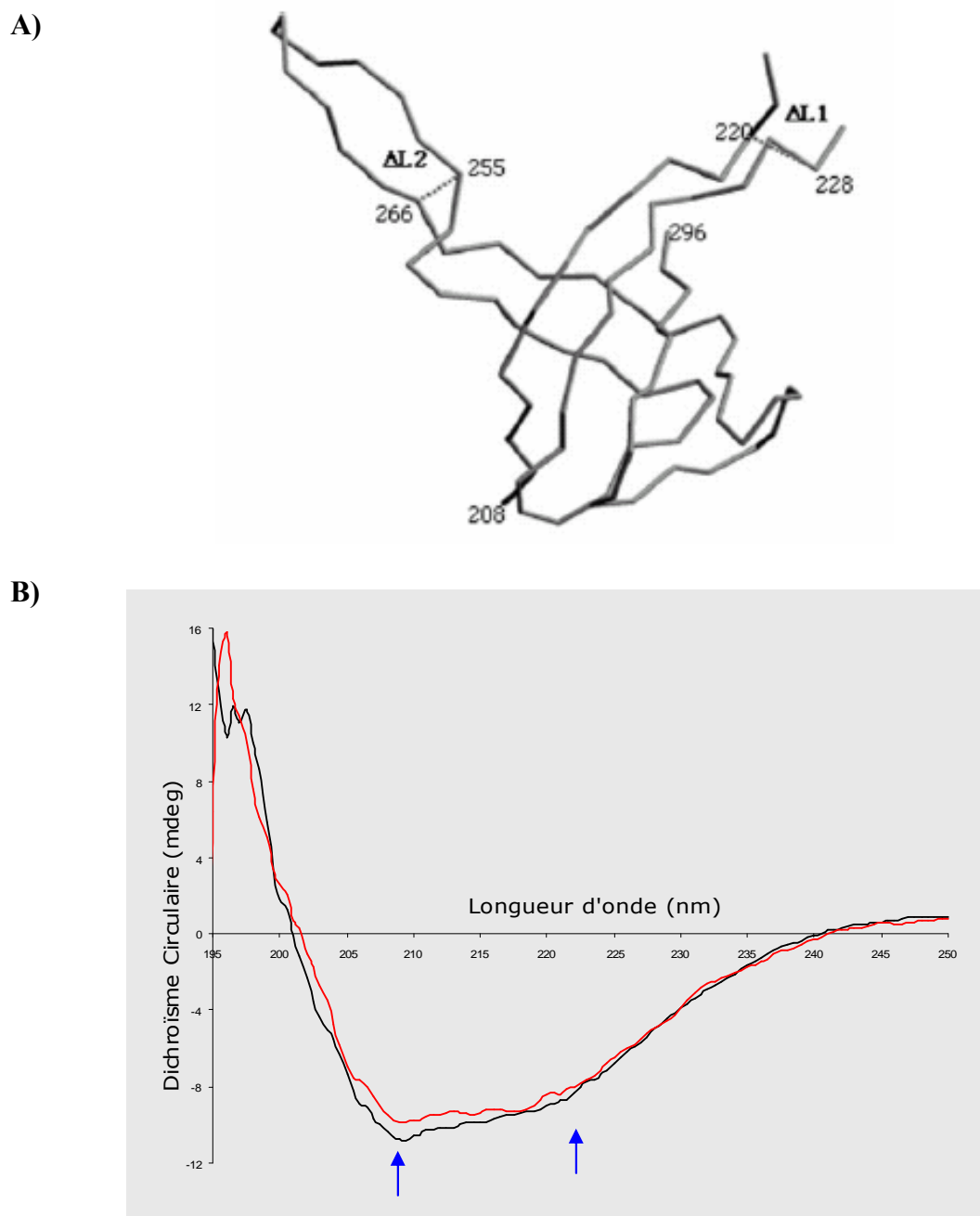
Ainsi, un seul des mutants envisagés présente un défaut d'association : le mutant  $\gamma_{\Delta L1}$  qui ne lie plus la sous-unité  $\alpha$ . Cette dernière pourrait donc interagir avec la boucle L1 du domaine 2 de Pa-aIF2 $\gamma$ .

Aucune zone d'interaction avec la sous-unité  $\beta$  n'a pu être identifiée par cette méthode malgré un balayage relativement étendu sur la surface de la protéine Pa-aIF2 $\gamma$ . Cependant,  $\beta$  pourrait posséder plusieurs points d'ancrage sur la sous-unité  $\gamma$  et dans ce cas, la rupture de l'interaction avec un seul d'entre eux peut ne pas être suffisante pour abolir totalement la liaison entre les deux sous-unités. Des mutants multiples de Pa-aIF2 $\gamma$  restent à tester pour vérifier cette hypothèse.

### **3) Vérification de la stabilité structurale du mutant $\gamma_{\Delta L1}$ .**

Les résultats obtenus par les tests d'assemblage sur gels natifs témoignent d'un rôle de la boucle L1 de  $\gamma$  dans la liaison à  $\alpha$  par l'intermédiaire d'un défaut de liaison du mutant  $\gamma_{\Delta L1}$ . Cependant, ce défaut de liaison pourrait être dû à un désordonnement local de la sous-unité  $\gamma$  qui affecterait sa liaison avec  $\alpha$  sans impliquer particulièrement la boucle L1. Un mauvais repliement de la protéine dans son ensemble est à exclure. En effet, le mutant  $\gamma_{\Delta L1}$  est toujours capable de lier la sous-unité  $\beta$ . Ce qui indique que la zone de fixation de  $\beta$  tout au moins présente un repliement correct. Par contre, une désorganisation locale de  $\gamma$  au niveau du domaine 2 ne peut pas être exclue sans plus d'arguments.

Pour vérifier que le mutant  $\gamma_{\Delta L1}$  acquiert la même structure globale et locale que la protéine native, nous avons réalisé des expériences comparatives de dichroïsme circulaire sur la protéine native et sur le mutant de la boucle L1. Le dichroïsme circulaire est une technique qui permet de mesurer, pour une molécule chirale, la différence d'absorption entre la composante polarisée circulairement gauche et la composante polarisée circulairement droite d'une lumière de polarisation plane qui traverse la molécule considérée. Dans le cas d'une protéine, si l'on se place sur une gamme de longueurs d'onde bien choisie (ici, les UV lointains), l'absorption sera caractéristique de la structure secondaire de la molécule. Ainsi, le spectre obtenu en réalisant un balayage de la gamme de longueur d'onde sélectionnée traduit l'arrangement structural de la protéine du point de vue de sa structure secondaire. Cette technique est par conséquent sensible à un désordonnement structural de la protéine, même si celui-ci reste local.



**Figure 3.20. : Mesure des spectres de dichroïsme circulaire de  $\gamma$  native et de  $\gamma_{\Delta L1}$  de *P. abyssii*.** **A)** localisation des délétions  $\Delta L1$  et  $\Delta L2$  réalisées dans le domaine 2 d'Pa- $\alpha$ IF2 $\gamma$ . Seul le tracé de la chaîne polypeptidique entre les différents C $\alpha$  de la molécule est dessiné. Les deux boucles pointent clairement hors du domaine et ne sont pas impliquées dans des interactions avec le cœur du tonneau  $\beta$ . La mutation  $\Delta L1$  correspond au remplacement des résidus 221 à 227 par une glycine. La distance entre les C $\alpha$  des deux résidus N220 et K228 est de 7.6 Å. Cette distance est compatible avec la longueur d'un segment comprenant 3 atomes C $\alpha$  dans une conformation étendue (6.5 Å classiquement). Ainsi, la liaison de la glycine à la place de la boucle ne devrait pas engendrer de distorsion dans la structure du tonneau  $\beta$ . **B)** Spectres de dichroïsme circulaire enregistrés entre 195 et 250 nm pour Pa- $\alpha$ IF2 $\gamma$  native (courbe noire) et pour Pa- $\alpha$ IF2 $\gamma_{\Delta L1}$  (courbe rouge). Les flèches bleues indiquent les deux minima caractéristiques à 208 (feuillets  $\beta$ ) et à 222 nm (hélices  $\alpha$ ).

Sur une gamme de longueurs d'onde allant de 190 à 250 nm, des minima d'absorption caractéristiques des deux types de structures secondaires d'une protéine classique sont observés :

- un minimum à 222 nm pour les structures en hélices  $\alpha$ . On retrouve également un minimum de plus petite amplitude à 209 nm et un maximum à 192 nm.
- un minimum significatif vers 208-210 nm pour les structures en feuillets  $\beta$ .

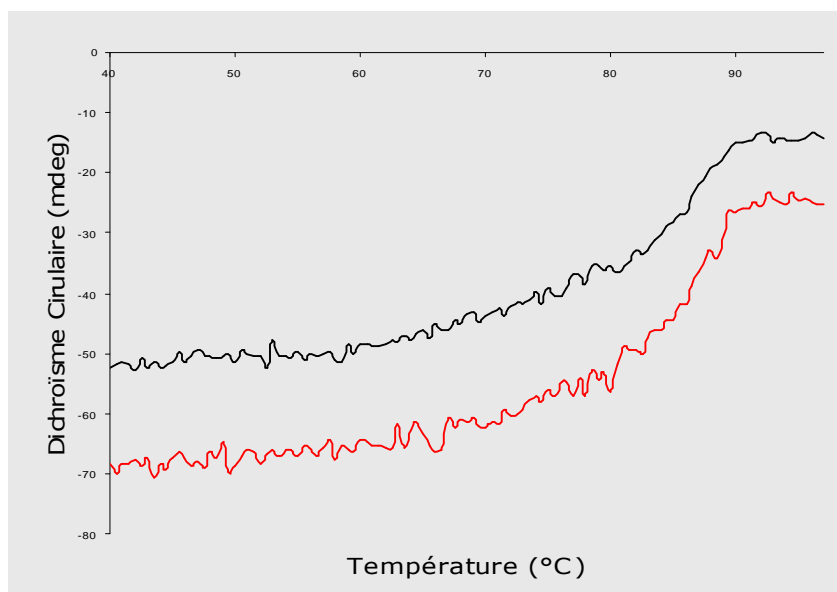
Des spectres de dichroïsme circulaire ont été enregistrés pour les deux protéines  $\gamma$  et  $\gamma_{\Delta L1}$  à l'Institut Curie à Orsay (voir la partie **Matériels et Méthodes** pour le protocole exact). Les protéines ont été pour cela diluées à une concentration de l'ordre de 2.5  $\mu\text{M}$  dans un tampon 10 mM phosphate de potassium pH 6.7, 100 mM NaCl. Les spectres ont été enregistrés à 20°C entre 195 et 250 nm. La gamme de longueur a été balayée trois fois et chaque spectre correspond à la moyenne de ces trois balayages successifs.

La **Figure 3.20** représente les spectres obtenus respectivement pour la protéine  $\gamma$  native et pour le mutant  $\gamma_{\Delta L1}$ . Ces deux spectres sont très similaires, avec un minimum très prononcé vers 208 nm. En complément de l'enregistrement de ces deux spectres, la dénaturation thermique irréversible des deux protéines a été suivie en mesurant leur absorption à 220 nm entre 40 et 95°C pour des concentrations protéiques de l'ordre de 15  $\mu\text{M}$  dans un tampon 10 mM phosphate de potassium pH 6.7, 100 mM NaCl. Les protéines  $\gamma$  et  $\gamma_{\Delta L1}$  présentent des courbes de dénaturation similaires (voir **Figure 3.21**) et dans les deux cas, une température de dénaturation  $T_m$  de 85°C est mesurée.

Le comportement similaire du mutant  $\gamma_{\Delta L1}$  à celui de la protéine  $\gamma$  native suggère que la mutation  $\Delta L1$  ne détruit pas la structure locale du domaine 2 de  $\gamma$ . Par conséquent, c'est bien l'omission de la boucle L1, et non la déstructuration que pourrait engendrer cette délétion, qui crée le défaut de liaison de la sous-unité  $\alpha$ . Ainsi, cet élément spécifique du facteur d'initiation e/aIF2 par rapport aux facteurs d'élongation est très certainement impliqué dans la fixation de Pa-aIF2 $\alpha$  sur Pa-aIF2 $\gamma$ .

Une étude réalisée sur aIF2 $\gamma$  de *Methanococcus jannaschii* est parue concomitamment à la publication de nos résultats (Roll-Mecak *et al.*, 2003). Ces travaux montrent que les deux mutants L256D et D325A d'aIF2 $\gamma$  de *M. jannaschii* présentent un très fort affaiblissement de leur capacité à lier la sous-unité  $\alpha$ . Chez Pa-aIF2 $\gamma$ , ces deux résidus correspondent respectivement à la leucine 229, qui termine la boucle L1 du domaine 2, et à l'aspartate 298. Le résidu D298 se trouve à la surface du domaine 2 de Pa-aIF2 $\gamma$ , dans le voisinage de la boucle L1. Ainsi, la zone que nous avons identifiée chez Pa-aIF2 $\gamma$  comme responsable de la

liaison à  $\alpha$  semble jouer le même rôle chez aIF2 $\gamma$  de *M. jannaschii*. Ces résultats viennent donc corroborer nos conclusions concernant le rôle de ces résidus dans la liaison à  $\alpha$ .



**Figure 3.21 :** Dénaturation thermique de Pa-aIF2 $\gamma$  native et de Pa-aIF2 $\gamma_{\Delta L1}$ . Spectres de dichroïsme circulaire enregistrés à 220 nm de 40 à 95°C pour Pa-aIF2 $\gamma$  native (courbe noire) et pour Pa-aIF2 $\gamma_{\Delta L1}$  (courbe rouge).

#### **4) Validation du modèle de liaison de l'ARNt fourni par EF1A.**

Comme cela a été décrit dans le **Chapitre II**, l'homologie de Pa-aIF2 $\gamma$  et d'EF1A a permis de réaliser un modèle d'amarrage de la molécule de Phe-ARNt<sup>Phe</sup> sur la sous-unité  $\gamma$  (**Figure 2.8**).

Les zones d'interaction potentielles qui sont mises en avant par ce modèle de positionnement se situent dans les domaines 2 et 3 de Pa-aIF2 $\gamma$ . De façon très intéressante, les régions spécifiques de Pa-aIF2 $\gamma$  que nous avons identifiées pour un éventuel rôle dans la liaison à  $\alpha$  coïncident relativement bien avec ces régions. Ainsi, la boucle L1 relie les brins  $\beta 7$  et  $\beta 8$  qui pourraient former la poche de liaison de la tige acceptrice de l'ARNt. De même, dans le domaine 3, la boucle L4 est très proche du bras T $\Psi$ C de l'ARNt. Bien qu'il s'agisse d'un élément caractéristique de  $\gamma$  par rapport à EF1A, cette boucle n'est pas totalement absente chez EF1A, elle est simplement plus courte que dans le cas de Pa-aIF2 $\gamma$ . Ces observations nous ont conduit à tester la capacité des différents mutants de  $\gamma$  à lier l'ARNt initiateur méthionylé. Le **Tableau 3.22** reprend les résultats obtenus en test de protection contre la déacylation avec ces différents mutants.

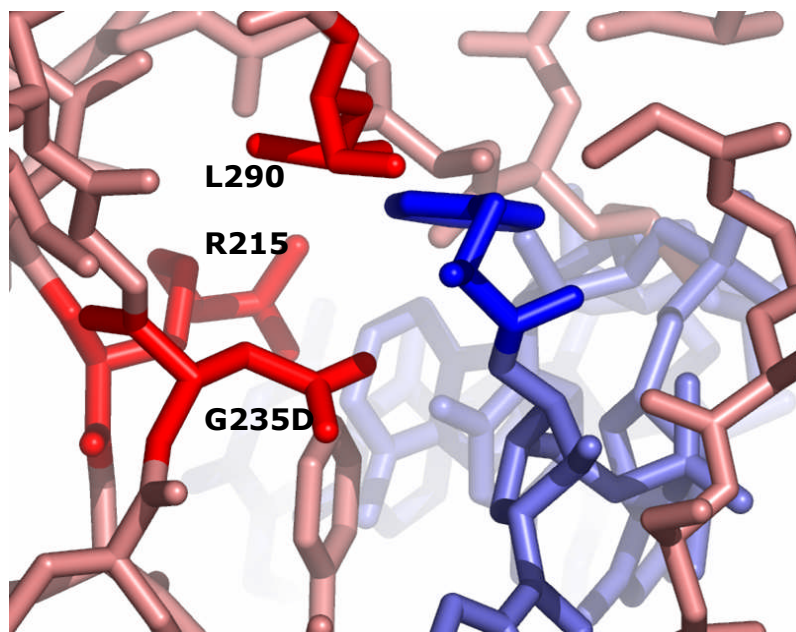
Ainsi, pour les mutants  $\gamma_{\Delta L3}$  et  $\gamma_{\Delta L4}$ , aucune protection contre la déacylation n'est observée par rapport à la sous-unité native qui présente un Kd de 5  $\mu$ M. Ces résultats ont été reproduits avec les hétérotrimères correspondants purifiés : aucune protection n'est réalisée par  $\alpha\beta\gamma_{\Delta L3}$  et  $\alpha\beta\gamma_{\Delta L4}$ . Pour le mutant  $\gamma_{\Delta L1}$ , qui ne lie plus  $\alpha$ , nous avons regardé si la présence de l'ARNt permettait de rétablir la liaison à  $\alpha$ . Or aucune protection n'est observée jusqu'à une concentration de 50  $\mu$ M pour chaque protéine. Ceci suggère d'une part qu'il n'y a pas de phénomène coopératif entre  $\gamma_{\Delta L1}$  et l'ARNt pour rétablir la liaison à  $\alpha$ , et d'autre part, que le mutant doit également présenter un défaut de liaison de l'ARNt, ce qui est confirmé en utilisant le monomère mutant seul. Contrairement à ces trois mutants pour lesquels la liaison avec l'ARNt est très fortement diminuée, le mutant  $\gamma_{\Delta L2}$  ne présente aucun défaut de liaison du Met-ARNt<sup>Met</sup> (Kd = 130 nM pour l'hétérotrimère  $\alpha\beta\gamma_{\Delta L2}$ , **Tableau 3.22**).

	$\alpha$	$\beta$	$\gamma$	Kd (nM)
1	natif	natif	natif	120 $\pm$ 40
2	natif		$\gamma_{\Delta L1}$	> 50 000
3	natif	natif	$\gamma_{\Delta L2}$	130 $\pm$ 60
4	natif	natif	$\gamma_{\Delta L3}$	> 5 000
5	natif	natif	$\gamma_{\Delta L4}$	> 5 000
6	natif	natif	$\gamma_{G235D}$	> 50 000
7			natif	5 000 $\pm$ 3 000
8			$\gamma_{\Delta L1}$	> 50 000
9			$\gamma_{\Delta L3}$	> 20 000
10			$\gamma_{\Delta L4}$	> 20 000

**Tableau 3.22 :** Constantes de dissociation des complexes réalisés entre le Met-ARNt<sup>Met</sup> d'*E. coli* à une concentration de 80 nM et les différents mutants d'Pa-aIF2 $\gamma$  de *P. abyssi*, sous forme trimérique ou monomérique.

Enfin, un dernier mutant a été testé, il s'agit du mutant  $\gamma_{G235D}$ . Ce mutant a été obtenu spontanément lors de la production de la sous-unité  $\gamma$  (Schmitt *et al.*, 2002). Ce résidu G235 correspond à la T239 d'EF1A qui se situe dans la poche de liaison de l'acide aminé et de la base 76 de l'ARNt. Or au sein de l'hétérotrimère  $\alpha\beta\gamma_{G235D}$ , la mutation annihile la liaison à l'ARNt puisque aucune protection n'est observée pour une concentration allant jusqu'à 50  $\mu$ M en protéine. La mutation G235D affecte la troisième glycine de la séquence consensus <sup>231</sup>GGΦΦGG<sup>236</sup> conservée chez tous les facteurs e/aIF2. Dans la structure d'aIF2 $\gamma$ -G235D de *P. abyssi*, le résidu aspartate mutant pointe à l'extérieur du tonneau  $\beta$ . Sa chaîne latérale pourrait ainsi gêner l'insertion de l'extrémité aminoacylée de la tige acceptrice de l'ARNt

dans la poche de liaison formée par les brins  $\beta 7$ ,  $\beta 8$  et  $\beta 14$  du domaine 2 de  $\gamma$ . Cet encombrement stérique pourrait ainsi expliquer l'absence de protection observée pour le mutant  $\gamma_{G235D}$  (voir **Figure 3.23**).



**Figure 3.23 :** détail de la tige acceptrice de l'ARNt dans la poche de liaison du domaine 2 de Pa-aIF2 $\gamma$  au sein du modèle de positionnement réalisé à partir de la structure du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup>. Figure dessinée avec le logiciel *PyMOL*. Le Phe-ARNt<sup>Phe</sup> positionné sur Pa-aIF2 $\gamma$  est représenté en bleu avec le résidu phénylalanine en bleu foncé. L'incompatibilité stérique observée dans ce modèle est visible sur cette figure, avec la leucine 290 qui se retrouve très proche du résidu Phe. La glycine 235 se trouve au fond de la poche de liaison. La chaîne de l'aspartate mutant pointe ici très nettement vers la partie aminoacide de l'ARNt. Enfin, à l'arrière-plan, l'arginine 215 est dirigée vers la tige acceptrice de l'ARNt.

Ces résultats permettent donc de valider le modèle de positionnement de l'ARNt sur Pa-aIF2 $\gamma$ . Ainsi, si l'on affecte la poche de liaison du domaine 2 (boucle L1 qui lie  $\beta 7$  et  $\beta 8$  ou résidu G235) ou la boucle L4 du domaine 3, la fixation de l'ARNt initiateur sur Pa-aIF2 $\gamma$  est très affaiblie. Ces résultats ont d'ailleurs été retrouvés chez eIF2 $\gamma$  de levure (Roll-Mecak *et al.*, 2003). En effet, la mutation du résidu G397 (équivalent à G292 du brin  $\beta 14$  chez Pa-aIF2 $\gamma$ ) en alanine entraîne chez la levure un défaut de croissance qui peut être contourné en surexprimant le Met-ARNt<sub>i</sub><sup>Met</sup>. La mutation G397I est quant à elle létale chez la levure. Cette glycine se situe à la base de la poche de fixation de la tige acceptrice de l'ARNt. La mutation de cette glycine en un résidu plus volumineux empêche ainsi la fixation de l'ARNt très probablement en provoquant une gêne stérique avec le ribose de la base 76 de l'ARNt. De même, la modification du R319 d'eIF2 $\gamma$  de levure en alanine conduit à un défaut de croissance. Ce résidu correspond à R215 de la séquence RSFD du brin  $\beta 7$  de Pa-aIF2 $\gamma$ . Ce

résultat confirme ainsi la participation de cette séquence à la liaison de l'ARNt. L'ensemble de ces données permet de valider le modèle de fixation de l'ARNt par Pa-aIF2 $\gamma$  selon un mode similaire à celui observé pour EF1A. La nature divergente des résidus impliqués dans la reconnaissance de l'ARNt chez les deux facteurs pourrait d'ailleurs expliquer la spécificité d'e/aIF2 $\gamma$  vis-à-vis de l'ARNt initiateur.

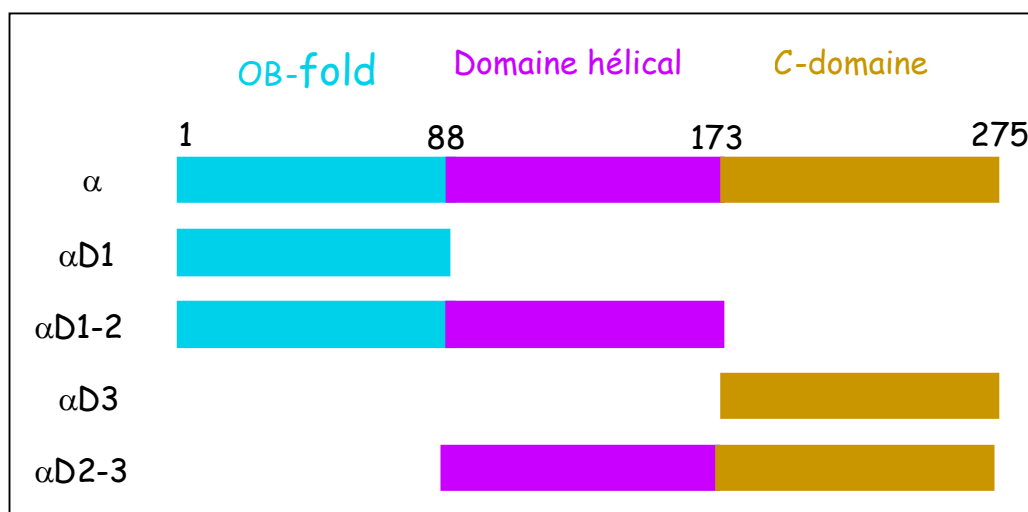
## **Partie C : Cartographie des zones d'interactions entre les sous-unités du facteur Pa-aIF2.**

### **La sous-unité $\alpha$ , un partenaire essentiel.**

Au sein de l'hétérotrimère, aIF2 $\alpha$  joue un rôle primordial pour l'obtention d'une affinité optimale vis-à-vis de l'ARNt. Une seconde partie de notre travail a consisté à identifier les zones de Pa-aIF2 $\alpha$  impliquées dans la liaison à  $\gamma$ . Dans le même temps, nous avons produit les domaines séparés de la sous-unité  $\beta$  pour vérifier que la zone d'interaction avec  $\gamma$  se trouve dans le domaine 1 de  $\beta$  (Hashimoto *et al.*, 2002).

#### **1) Production des domaines séparés de Pa-aIF2 $\alpha$ .**

Comme indiqué dans le **Chapitre II**, la sous-unité Pa-aIF2 $\alpha$  est organisée en trois domaines et la structure des domaines 1 et 2 a été publiée au cours de ce travail de thèse pour les protéines eIF2 $\alpha$  humaine et de levure (Nonato *et al.*, 2002 ; Dhaliwal and Hoffman, 2003). Pour déterminer quelle(s) partie(s) de la protéine est(sont) impliquée(s) dans l'interaction avec la sous-unité  $\gamma$ , les domaines séparés d' $\alpha$  ont été construits. La **Figure 3.24** reprend les différentes constructions réalisées.



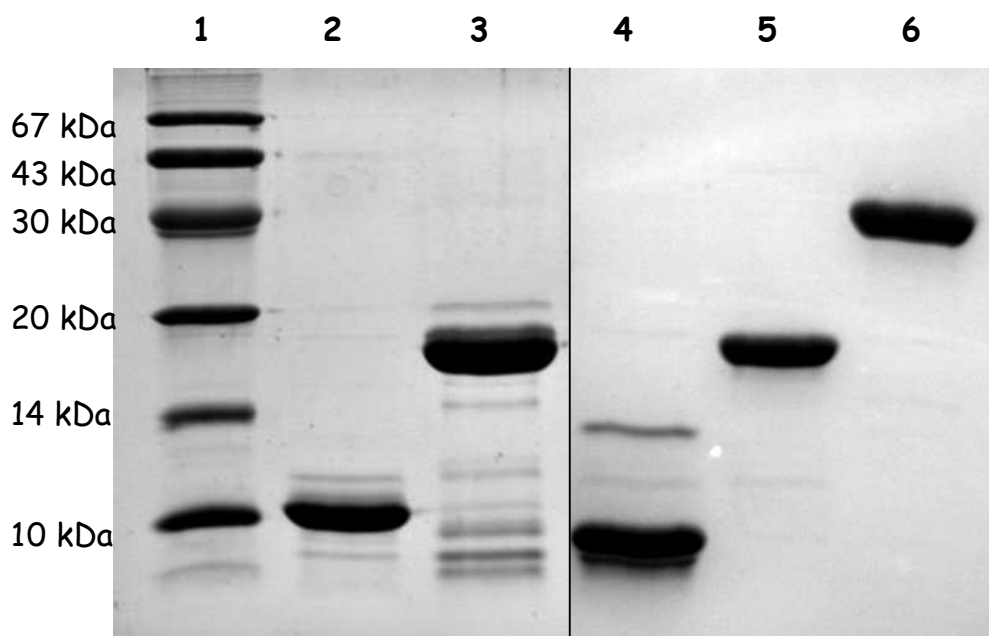
**Figure 3.24 :** production des domaines séparés d'aIF2 $\alpha$  de *P. abyssi*.

Les ADNs des formes  $\alpha_{D1}$  et  $\alpha_{D1-2}$  ont été obtenus par mutagenèse dirigée par PCR en utilisant la technique *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene. Les amorces oligonucléotidiques utilisées ont permis d'introduire un codon stop au niveau des résidus 88 et 173, respectivement. Les ADNs des formes  $\alpha_{D3}$  et  $\alpha_{D2-3}$  ont été obtenus par mutagenèse

dirigée par PCR de chevauchement en utilisant des amorces internes permettant d'introduire un codon de démarrage au niveau des résidus 173 et 88, respectivement. Les fragments amplifiés ont ensuite été reclés dans pET3alpa. Pour chaque forme, les plasmides obtenus ont été séquencés pour vérifier la conformité de la construction.

La même méthode a été utilisée pour produire les domaines 1 (résidus 1 à 93) et 2 (résidus 94 à 144) de la sous-unité Pa-aIF2 $\beta$ .

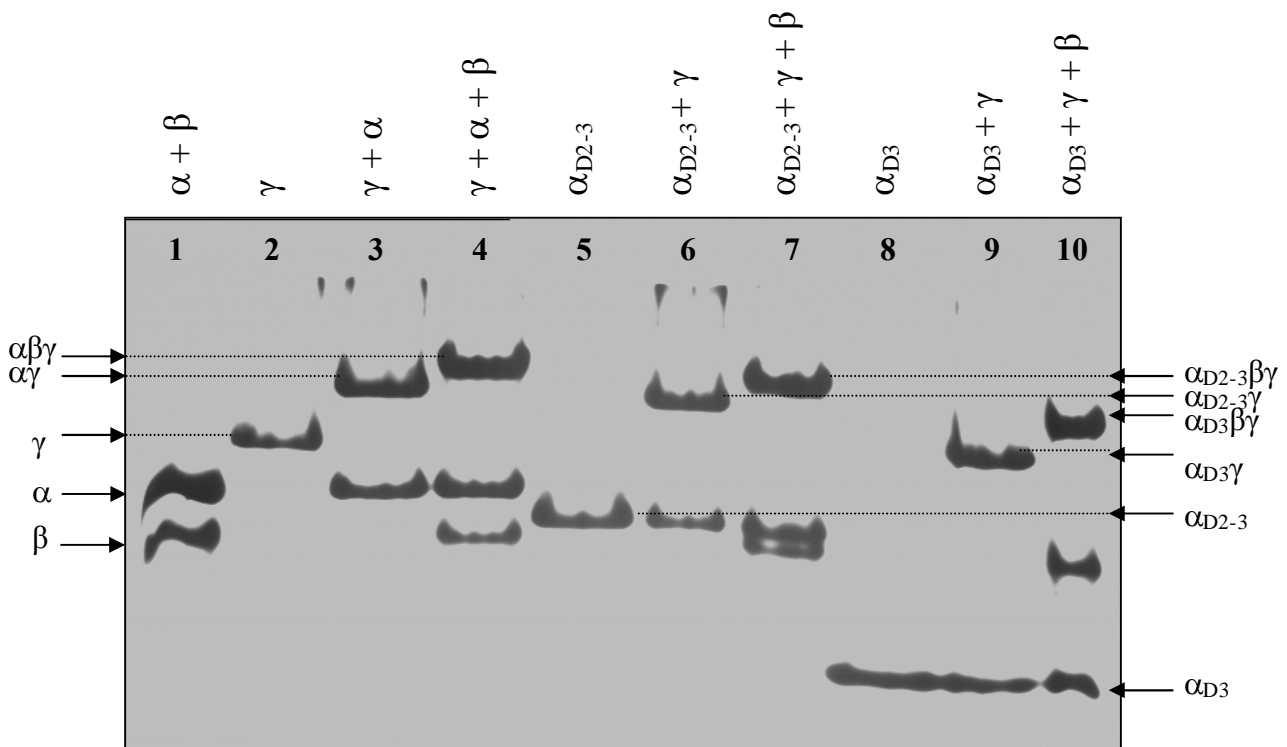
Les formes tronquées de Pa-aIF2 $\alpha$  et de Pa-aIF2 $\beta$  ont été exprimées de la même manière que les protéines natives. Les protéines Pa-aIF2 $\beta_1$  et Pa-aIF2 $\beta_2$  ont été purifiées suivant le même protocole que Pa-aIF2 $\beta$ . La purification des protéines Pa-aIF2 $\alpha_{D1}$  et Pa-aIF2 $\alpha_{D3}$  a été réalisée suivant la même méthode que Pa-aIF2 $\alpha$ . Cependant, l'étape de purification sur la colonne S-Sépharose a été effectuée à plus basse force ionique (100 mM NaCl). Enfin, les protéines Pa-aIF2 $\alpha_{D1-2}$  et Pa-aIF2 $\alpha_{D2-3}$  présentant un pI acide (6.5), l'étape de purification sur la S-Sepharose a été remplacée par un passage sur une colonne échangeuse d'anions de type Q-Hiload dans le même tampon. L'homogénéité des différentes protéines ainsi purifiées a été vérifiée sur gel de polyacrylamide en conditions dénaturantes (voir **Figure 3.25**).



**Figure 3.25 :** Production des formes tronquées de Pa-aIF2 $\alpha$ . Analyse des échantillons purifiés par migration électrophorétique sur gel de polyacrylamide 16% dénaturant. Piste 1 : marqueur de poids moléculaire ; Piste 2 :  $\alpha_{D3}$  ; Piste 3 :  $\alpha_{D2-3}$  ; Piste 4 :  $\alpha_{D1}$  ; Piste 5 :  $\alpha_{D1-2}$  ; Piste 6 :  $\alpha$  native.

## 2) Tests d'assemblages avec la sous-unité $\gamma$ .

La capacité de ces différentes formes tronquées d' $\alpha$  et de  $\beta$  à lier la sous-unité  $\gamma$  a ensuite été vérifiée par assemblage sur gel d'électrophorèse natif. Le protocole est identique à celui utilisé pour tester les mutants de Pa-aIF2 $\gamma$ .



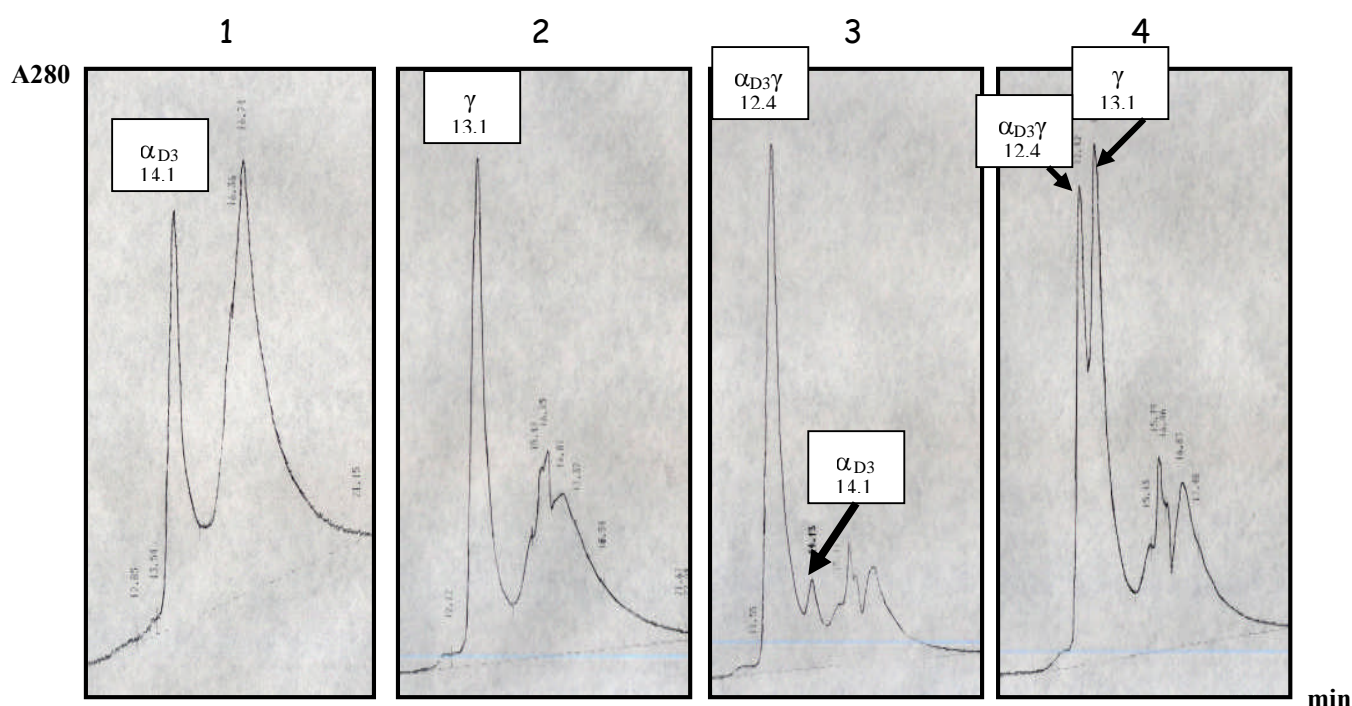
**Figure 3.26 : Assemblage des formes tronquées  $\alpha_{D2-3}$  et  $\alpha_{D3}$  de Pa-aIF2 $\alpha$  avec la sous-unité  $\gamma$ .** Les protéines indiquées au-dessus de chaque piste ont été préalablement mélangées dans 10  $\mu$ L de tampon 200 mM NaCl, 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol. 10  $\mu$ L de tampon de dépôt (240 mM KOH, 144 mM acide acétique, 20% glycérol, 0.4% (w/v) bleu de méthylène) ont été ajoutés et les échantillons ont été déposés sur un gel natif acide (12% en polyacrylamide). Après migration, les gels ont été colorés au bleu de Coomassie. Piste 1 :  $\alpha + \beta$  (30  $\mu$ M et 50  $\mu$ M respectivement) ; Pistes 2 à 4 :  $\gamma$  (19  $\mu$ M, piste 2) mélangé avec un excès molaire d' $\alpha$  (piste 3), ou d' $\alpha + \beta$  (piste 4) ; Pistes 5 à 7 :  $\alpha_{D2-3}$  (35  $\mu$ M, piste 5),  $\gamma$  (19  $\mu$ M) mélangé avec un molaire d' $\alpha_{D2-3}$  (piste 6), ou d' $\alpha_{D2-3} + \beta$  (piste 7) ; Pistes 8 à 10 :  $\alpha_{D3}$  (50  $\mu$ M, piste 8),  $\gamma$  (19  $\mu$ M) mélangé avec un excès molaire d' $\alpha_{D3}$  (piste 9), ou d' $\alpha_{D3} + \beta$  (piste 10).

Les deux formes tronquées  $\alpha_{D2-3}$  et  $\alpha_{D3}$  de Pa-aIF2 $\alpha$  sont capables de lier Pa-aIF2 $\gamma$  (**Figure 3.26**). A l'instar des assemblages réalisés avec la sous-unité native, les deux hétérodimères et les deux hétérotrimères sont formés avec ces mutants (pistes 6 et 7 pour  $\alpha_{D2-3}$  et pistes 9 et 10 pour  $\alpha_{D3}$ ). La liaison a été confirmée en reproduisant ces gels natifs en se plaçant en excès de  $\gamma$  : toute la protéine  $\alpha_{D3}$  est consommée dans la formation de l'hétérodimère  $\alpha_{D3}\gamma$ . Par contre, le domaine 1 isolé d' $\alpha$  n'interagit pas avec la sous-unité  $\gamma$ . De même, aucune interaction n'est observée avec la forme correspondant aux domaines 1 et 2 (résultats obtenus sur gel d'électrophorèse de manière similaire). Ces différentes observations

nous permettent de conclure que le domaine 3 de Pa-aIF2 $\alpha$  est nécessaire et suffisant pour réaliser l'interaction avec Pa-aIF2 $\gamma$ .

Des expériences similaires avec les deux formes tronquées de  $\beta$  ont permis de confirmer que le site de fixation de  $\gamma$  sur  $\beta$  se trouvait dans le domaine  $\beta_1$ , le domaine 2 de  $\beta$  n'interagissant pas avec  $\gamma$ .

Ces expériences d'assemblage par électrophorèse native ont également été confirmées par des tests d'assemblages réalisés par chromatographie d'exclusion. La **Figure 3.27** illustre les types de chromatogrammes obtenus avec l'assemblage  $\alpha_{D3}\gamma$ .



**Figure 3.27 :** Test d'assemblage de la sous-unité  $\gamma$  avec la forme tronquée  $\alpha_{D3}$  d'Pa-aIF2 $\alpha$  de *P. abyssi* par chromatographie d'exclusion. Les chromatogrammes représentés correspondent aux injections suivantes : 1, 100  $\mu$ L d' $\alpha_{D3}$  17  $\mu$ M ; 2, 100  $\mu$ L de  $\gamma$  12  $\mu$ M ; 3, 100  $\mu$ L de  $\gamma$  12  $\mu$ M +  $\alpha_{D3}$  25  $\mu$ M ; 4, 100  $\mu$ L de  $\gamma$  17  $\mu$ M +  $\alpha_{D3}$  8  $\mu$ M. Les temps de rétention des différentes protéines sont indiqués dans l'encart.

Dans le cas de l'assemblage  $\alpha_{D3}\gamma$  représenté sur la **Figure 3.27**, la formation de l'hétérodimère est effective, que  $\gamma$  soit en défaut (chromatogramme n°3) ou en excès (chromatogramme n°4).

Le **Tableau 3.28** récapitule les nouveaux résultats obtenus pour Pa-aIF2 $\alpha$  et Pa-aIF2 $\beta$ .

	Forme Tronquée	Assemblage avec $\gamma$
1	$\alpha_{D1}$	-
2	$\alpha_{D1-2}$	-
3	$\alpha_{D2-3}$	+
4	$\alpha_{D3}$	+
5	$\beta_1$	+
6	$\beta_2$	-

**Tableau 3.28 :** Récapitulatif des test d’assemblages avec Pa-aIF2 $\gamma$  réalisés par chromatographie d’exclusion et par électrophorèse sur gel natif en présence des différentes formes tronquées de Pa-aIF2 $\alpha$  et de Pa-aIF2 $\beta$ .

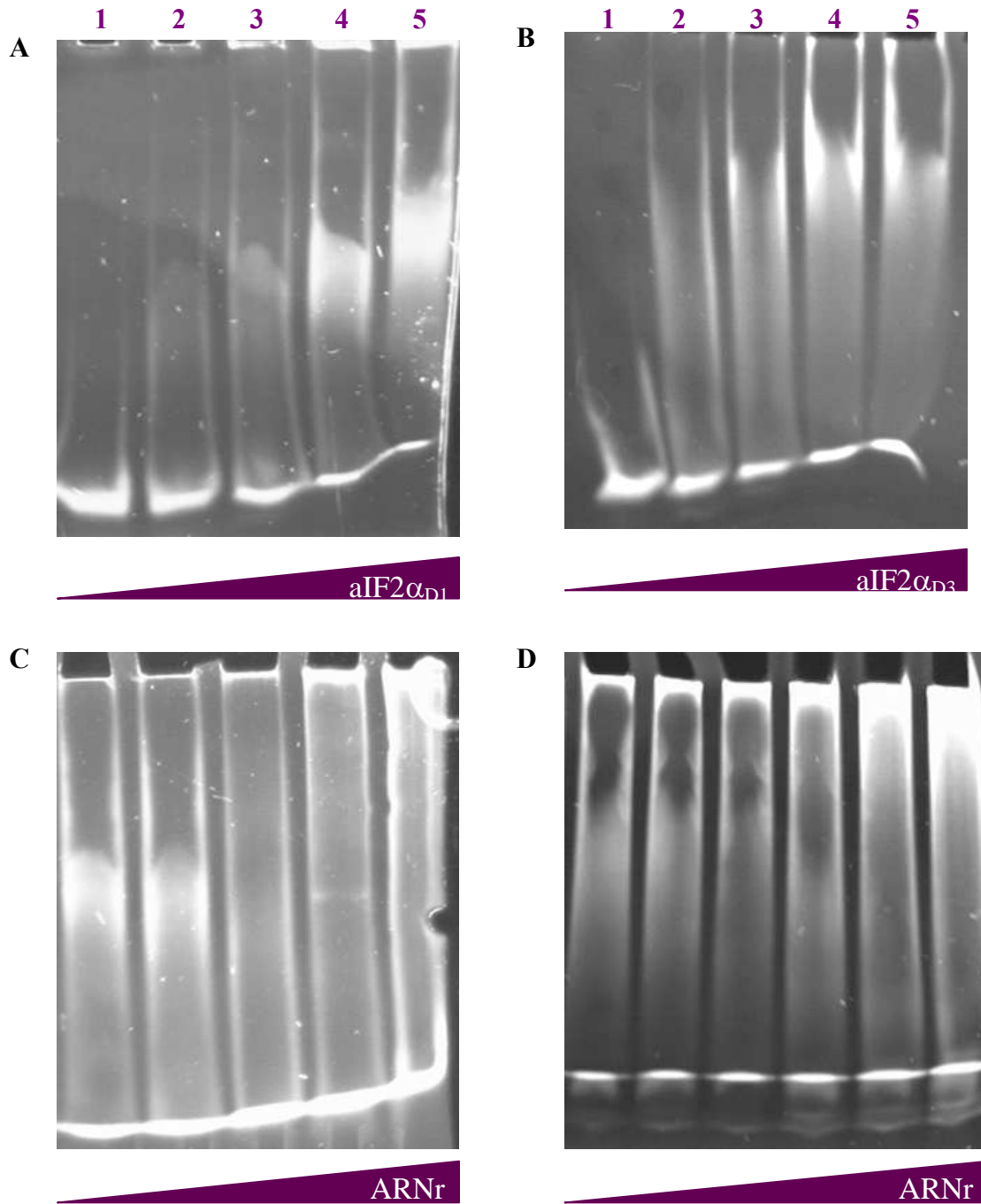
### 3) Implications fonctionnelles.

Les expériences précédentes montrent que Pa-aIF2 $\alpha_{D3}\gamma$  est un hétérodimère stable. Dans un second temps, nous avons étudié la capacité de ce mini-dimère à lier l’ARNt (Tableau 3.29).

	$\alpha$	$\beta$	$\gamma$	Kd (nM)
1	natif	natif	natif	120 $\pm$ 40
2	natif		natif	44 $\pm$ 20
3			natif	5 000 $\pm$ 3 000
4	$\alpha_{D3}$		natif	42 $\pm$ 20
5	$\alpha_{D1}$		natif	5 000 $\pm$ 3 000
6	$\alpha_{D1-2}$		natif	5 000 $\pm$ 3 000

**Tableau 3.29 :** Constantes de dissociation des complexes réalisés entre le Met-ARNt<sub>f</sub><sup>Met</sup> d’*E. coli* à une concentration de 80 nM et les différentes formes tronquées de Pa-aIF2 $\alpha$  associées à Pa-aIF2 $\gamma$ .

Les tests de protection contre la déacylation spontanée montrent que l’hétérodimère tronqué  $\alpha_{D3}\gamma$  présente une affinité pour l’ARNt initiateur identique à celle de l’hétérodimère natif. Le domaine 3 de Pa-aIF2 $\alpha$  agit donc en coopération avec la sous-unité  $\gamma$  pour conférer à cette dernière une capacité de liaison de l’ARNt optimale. Ces résultats confirment de plus l’absence de rôle joué par  $\beta$  dans la liaison de l’ARNt. La question que l’on peut alors se poser est de savoir si  $\alpha_{D3}$  renforce l’affinité de  $\gamma$  pour l’ARNt par une interaction directe avec le polyribonucléotide ou bien si sa contribution est indirecte.



**Figure 3.30 :** Expériences de retard de migration sur gel de l'ARNt<sup>Met</sup> d'*E. coli* en présence des domaines 1 ou 3 d'Pa-aIF2 $\alpha$  de *P. abyssi*. **A)** 1.8  $\mu$ M d'ARNt<sup>Met</sup> d'*E. coli* ont été incubés dans 20  $\mu$ L de tampon 200 mM NaCl, 50 mM TBE pH 8.3, 10 mM 2-mercaptoéthanol, 1 mM MgCl<sub>2</sub>, 1 mM GDPNP en présence de concentrations croissantes d'Pa-aIF2 $\alpha$ <sub>D1</sub> (0, 10, 20, 40 et 80  $\mu$ M, de la piste 1 à la piste 5) de *P. abyssi*. Après 15 minutes d'incubation à température ambiante, les échantillons sont chargés sur un gel à 5% en polyacrylamide contenant 50 mM TBE pH 8.3 et 1 mM MgCl<sub>2</sub>. La migration a lieu dans le même tampon. Le gel est ensuite révélé par coloration au BET. **B)** Même expérience réalisée en présence d'Pa-aIF2 $\alpha$ <sub>D3</sub> cette fois-ci. **C)** Expériences de compétition entre l'ARNt et un mélange d'ARN ribosomiaux 16S + 23S pour la liaison aux domaines d' $\alpha$ . 80  $\mu$ M d'Pa-aIF2 $\alpha$ <sub>D1</sub> de *P. abyssi* et 1  $\mu$ g (1.8  $\mu$ M) d'ARNt<sup>Met</sup> d'*E. coli* ont été incubés dans 20  $\mu$ L de tampon 200 mM NaCl, 50 mM TBE pH 8.3, 10 mM 2-mercaptoéthanol, 1 mM MgCl<sub>2</sub>, 1 mM GDPNP. Des quantités croissantes du mélange d'ARNr 16S et 23S d'*E. coli* (1, 2, 4, 7.5 et 10  $\mu$ g de la Piste 1 à la Piste 5, respectivement) ont été ajoutées dans le mélange. **D)** Même expérience avec d'Pa-aIF2 $\alpha$ <sub>D3</sub> (80  $\mu$ M). Les quantités d'ARNr utilisées sont respectivement 0.5, 1, 2, 3, 5 et 7.5 de la piste 1 à la piste 6.

Pour tenter de répondre à cette question, nous avons étudié la capacité de liaison aux acides nucléiques des différents domaines d' $\alpha$ . Pour cela, nous avons effectué des expériences de retard de migration de l'ARNt sur gel d'électrophorèse en utilisant les formes tronquées de Pa-aIF2 $\alpha$  (**Figure 3.30**).

Le domaine 1 de Pa-aIF2 $\alpha$  présente une structure de type OB-fold, qui est un motif fréquemment impliqué dans la liaison des acides nucléiques (Murzin, 1993). Les expériences de retard de migration sur gel de l'ARNt montrent que le domaine 1 d' $\alpha$  est effectivement capable de lier l'ARNt. Par contre, la capacité de liaison d' $\alpha_{D1}$  n'est pas spécifique de ce motif ARNt, puisque le mélange ARN ribosomal 16S/23S entre en compétition avec l'ARNt pour la liaison à  $\alpha_{D1}$ . D'autre part, nous montrons que le domaine 3 de Pa-aIF2 $\alpha$  est également capable de lier l'ARNt (**Figure 3.30 B**). Par contre, là encore, cette interaction n'est pas spécifique du motif ARNt puisque l'ARN ribosomal est un bon compétiteur de cette interaction.

### **Conclusions au Chapitre III.**

Ainsi, le domaine C-terminal de Pa-aIF2 $\alpha$  est suffisant pour conférer à la sous-unité  $\gamma$  une capacité de liaison optimale de l'ARNt. Ce domaine se lie à aIF2 $\gamma$  *via* la boucle L1 du domaine 2 de  $\gamma$ . De plus, il possède des propriétés générales de liaison des acides nucléiques qui suggèrent qu'il pourrait contribuer à un renforcement de l'affinité pour l'ARNt par une interaction directe avec le polynucléotide. Cependant, cette affinité pour les ARNs n'est pas spécifique de l'ARNt. Par conséquent, on ne peut exclure l'hypothèse qu' $\alpha$  intervient seulement de manière indirecte en induisant un ajustement conformationnel de la sous-unité  $\gamma$  qui permette une liaison optimale de l'ARNt initiateur. Enfin, les propriétés de liaison des ARNs observées pour le domaine 1 d' $\alpha$  évoquent la possibilité pour ce domaine de communiquer avec un autre partenaire de type ARN intervenant dans le démarrage de la traduction et qui pourrait être l'ARNm ou un ARN ribosomal.

Ces résultats ont fait l'objet d'une publication dans *the Journal of Biological Chemistry* (**voir Liste des Travaux publiés**).

Pendant ce travail de thèse, deux structures d'un fragment de la protéine eucaryotique correspondant aux domaines 1 et 2 d' $\alpha$  ont été obtenues. Il s'agit de la structure des domaines 1 et 2 d'eIF2 $\alpha$  humaine (Nonato *et al.*, 2002) et de celle des mêmes domaines d'eIF2 $\alpha$  de levure (Dhaliwal and Hoffman, 2003). Dans le contexte des résultats acquis pour le rôle du domaine 3, la détermination de la structure tridimensionnelle de la sous-unité  $\alpha$  entière ou du domaine 3 isolé devient particulièrement intéressante.



*Chapitre IV : Etude  
Structurale de la  
Sous-Unité  $\alpha$ IF2 $\alpha$*



## **Partie A : Résolution de la structure de la sous-unité aIF2 $\alpha$ de l'archée *Pyrococcus abyssi*.**

Dans cette partie, nous avons cherché à résoudre la structure du domaine 3 de Pa-aIF2 $\alpha$  et celle de la sous-unité entière. Nous nous sommes dans un premier temps intéressés à la forme tronquée Pa-aIF2 $\alpha_{D2-3}$ .

### **1) Principe de la cristallographie par diffraction des rayons X.**

Lorsqu'un rayonnement électromagnétique, et en particulier les rayons X, traverse la matière, il interagit avec les particules rencontrées. Deux phénomènes résultent de cette interaction : une absorption et une diffusion du rayonnement incident dans toutes les directions de l'espace. Pour une onde électromagnétique, la diffusion est essentiellement due aux interactions avec les électrons. Lorsque la matière traversée présente une organisation spatiale périodique – comme c'est le cas pour un cristal – un phénomène particulier se produit : la diffraction. Les ondes diffusées par chacun des électrons du réseau cristallin interfèrent les unes avec les autres et le faisceau de rayons X ne diffracte alors que dans certaines régions privilégiées de l'espace.

Ainsi, si l'on observe le rayonnement diffusé dans une direction  $\vec{s}$  de l'espace, l'onde diffractée sera non nulle uniquement pour la famille de plans réticulaires (hkl) (où h, k et l sont les indices de Miller de la famille de plans) satisfaisant à la loi de Bragg :

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

où d est la distance entre deux plans (hkl) consécutifs du réseau

$\lambda$  la longueur d'onde du rayonnement incident

$\theta$  l'angle entre le rayonnement incident et le plan réticulaire (hkl)

n un entier positif.

A cette famille de plans réticulaires correspond donc un point dans l'espace de diffraction - espace appelé réseau réciproque. Le cliché de diffraction est alors constitué de tâches correspondant chacune à une famille de plans (hkl) et l'espace réciproque qu'il définit est donc directement lié à la densité électronique du cristal. Le facteur de structure  $F_{hkl}$  caractérise décrit l'onde diffractée en amplitude et en phase par les plans réticulaires (hkl).

$$F_{hkl} = \iiint_V \rho(xyz) e^{2i\pi(hx+ky+lz)} dx dy dz$$

où V est le volume de la maille et  $\rho$  la densité électronique.

Dans le cas de cristaux moléculaires, ce facteur peut s'écrire comme une somme discrète en fonction des facteurs de diffusion atomique  $f_j$  de chacun des N atomes de la maille  $(x_j, y_j, z_j)$  :

$$F_{hkl} = \sum_j f_j e^{2i\pi(hx_j+ky_j+lz_j)}$$

C'est un nombre complexe qui peut s'écrire sous forme d'un module et d'une phase:

$$F_{hkl} = |F_{hkl}| \times e^{i\phi(hkl)}$$

L'intensité diffractée  $I_{hkl}$  est proportionnelle au carré du module de  $F_{hkl}$  et s'obtient directement à partir des clichés de diffraction. Par conséquent, il reste à mesurer la phase  $\phi(hkl)$  pour chaque réflexion. La densité électronique en chaque point du cristal s'obtient alors par transformée de Fourier inverse, fonction qui prend la forme d'une somme finie sur tous les nœuds du réseau dans le cas d'un cristal :

$$\rho(xyz) = (1/V) \sum_h \sum_k \sum_l F_{hkl} e^{-2i\pi(hx+ky+lz)}$$

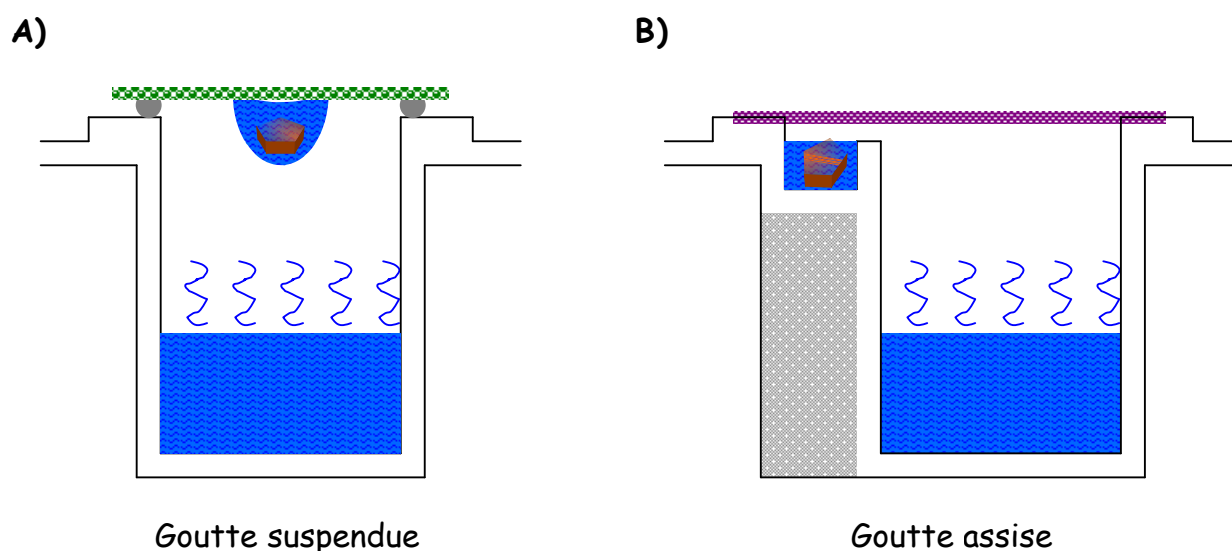
L'information relative aux phases étant perdue sur les clichés de diffraction, différentes méthodes ont été mises au point pour les calculer (MIR, MAD, remplacement moléculaire). La détermination des phases est l'étape limitante du processus de résolution des structures cristallographiques.

## **2) Obtention des cristaux d'aIF2 $\alpha_{D2-3}$ de P. abyssi.**

La cristallisation d'une macromolécule est un processus complexe. Elle peut se décomposer en quatre phases successives : formation d'un état de sursaturation, nucléation, croissance cristalline et cessation de croissance. Cependant, ces molécules biologiques présentent de nombreuses sources potentielles d'hétérogénéité qui peuvent nuire à leur capacité de cristalliser. On peut citer quelques exemples : la présence de contaminants qui n'ont pas été éliminés lors de la purification de la protéine, des microhétérogénéités de séquences, des formes de la protéine partiellement protéolysées ou ayant subi des modifications post-traductionnelles. Des régions flexibles au sein de la molécule peuvent

également perturber sa cristallisation. De plus, les macromolécules sont très sensibles à leur environnement (température, pH...) et elles peuvent se dégrader facilement.

La plupart des méthodes de cristallisation utilisent un agent précipitant qui permet de diminuer la solubilité de la protéine. Les principaux composés utilisés sont les sels (sulfate d'ammonium, NaCl,...), les polymères (le PEG est le plus couramment utilisé), les solvants organiques (2-méthylpentanediol ou MPD) et les composés organiques non volatils. Il existe différentes techniques permettant la mise en œuvre d'un processus de cristallisation : la cristallisation en « batch », la diffusion de liquide (par dialyse) et la diffusion de vapeur. La diffusion de vapeur est la technique la plus couramment utilisée au laboratoire. Dans cette technique, la sursaturation est atteinte par évaporation d'espèces volatiles. Elle consiste à placer, dans une enceinte étanche, une goutte (1 à 20  $\mu$ L) d'un mélange contenant la protéine à cristalliser, l'agent précipitant et un tampon, à proximité d'un réservoir de plus grand volume (100 à 500  $\mu$ L) contenant la solution d'agent précipitant. Les molécules volatiles vont alors diffuser entre le réservoir et la goutte et permettre éventuellement à la protéine d'atteindre progressivement l'état de sursaturation, condition indispensable à la nucléation. La **Figure 4.1** représente les deux principales techniques de la diffusion de vapeur utilisées au cours de ce travail : la technique de la goutte suspendue (obtention manuelle des cristaux) et la technique de la goutte assise (robot).

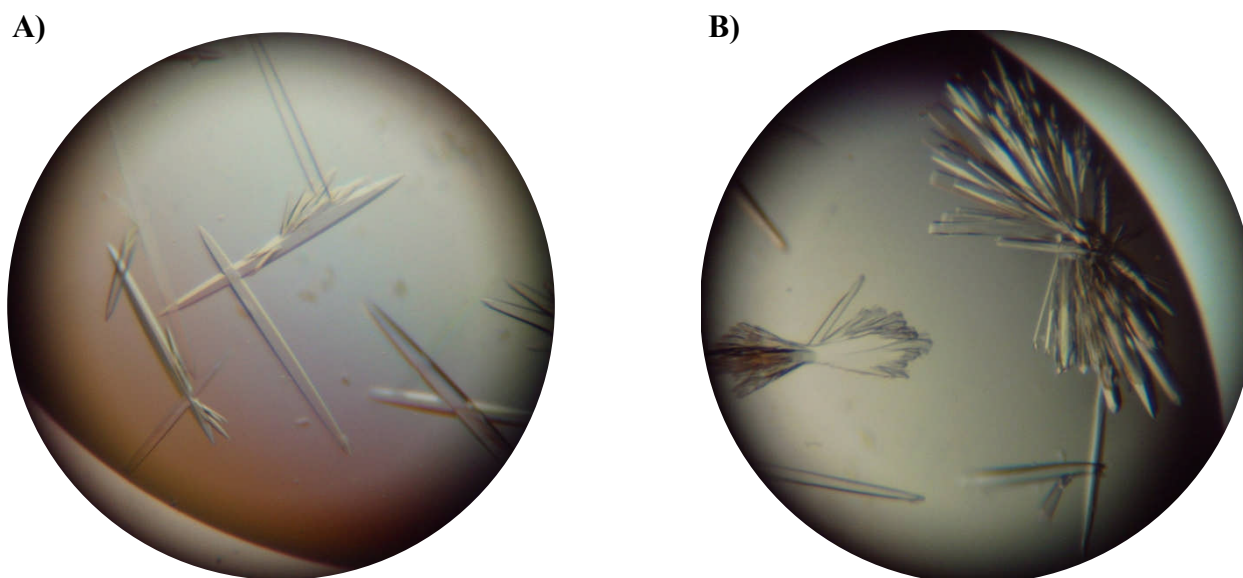


**Figure 4.1 :** Applications de la technique de la diffusion de vapeur. Pour la technique de la goutte suspendue **A**), la goutte (4  $\mu$ L) contenant la solution protéique et l'agent précipitant est déposée sur la lamelle de verre (en vert) avant d'être placée au-dessus du réservoir (500  $\mu$ L) contenant la solution d'agent précipitant. La graisse de silicone assure l'étanchéité de l'ensemble. La technique de la goutte assise **B**) mise en œuvre lors des essais de cristallisation à l'aide du robot repose sur la présence d'une cupule dans laquelle est déposée la goutte (1 à 2  $\mu$ L) au-dessus du réservoir (100  $\mu$ L). Un film plastique (en violet) referme le puits et assure ainsi l'étanchéité.

La cristallisation de la forme tronquée  $\alpha_{D2-3}$  de la sous-unité aIF2 $\alpha$  de *P. abyssi* a été réalisée à partir de la technique de la diffusion de vapeur en goutte assise à l'aide du robot *Genesis RSP100* (TECAN). Les matrices d'échantillonnage commerciales fournies par Hampton Research ont été utilisées pour la recherche des conditions de cristallisation de la protéine. Les premiers essais de cristallisation ont été effectués à 24°C à partir d'une solution de protéine  $\alpha_{D2-3}$  fraîchement purifiée suivant le protocole décrit et concentrée à 7 ou 14 mg.mL<sup>-1</sup>. Des gouttes de 2  $\mu$ L ont été réalisées en mélangeant 1  $\mu$ L de la solution protéique et 1  $\mu$ L de la solution contenant l'agent précipitant. Après une semaine, l'observation des gouttes a révélé la présence de cristaux ayant la forme de longues baguettes, ces dernières croissant sous forme d'amas étoilés. Ces cristaux ont été obtenus dans la condition de cristallisation n°5 de la matrice d'échantillonnage « Crystal Screen II » de Hampton (**Tableau 4.2**).

Condition	Agent Précipitant	Additif
CSII - 5	2.0 M Sulfate d'ammonium	5 % isopropanol

**Tableau 4.2 :** Condition de cristallisation de la forme aIF2 $\alpha_{D2-3}$  de *P. abyssi*.



**Figure 4.3 :** Cristaux de Pa-aIF2 $\alpha_{D2-3}$ . Les deux photos représentent les cristaux obtenus dans la condition CSII-5 (2.0 M sulfate d'ammonium, 5 % isopropanol). La photo **A**) montre les cristaux isolés en formes de longues baguettes fines et la photo **B**) illustre la croissance de ces cristaux sous forme d'amas étoilés.

Les cristaux ainsi obtenus (**Figure 4.3**) sont parfaitement reproductibles dans les mêmes conditions de cristallisation. Pour obtenir des cristaux de plus grande taille, la cristallisation de la protéine a été reproduite manuellement en boîtes de 24 puits avec des gouttes de 4  $\mu\text{L}$  (2  $\mu\text{L}$  de solution protéique + 2  $\mu\text{L}$  de solution du réservoir). Seules deux gouttes réalisées en présence de la même solution précipitante (2.0 M sulfate d'ammonium, 5 % isopropanol) ont donné naissance à des cristaux d' $\alpha_{D2-3}$  de taille légèrement supérieure à ceux obtenus avec le robot de cristallisation. La variation des concentrations de sulfate d'ammonium et d'isopropanol dans la solution précipitante n'a pas entraîné d'amélioration de la taille des cristaux.

### **3) Acquisition et traitement des données.**

#### **1. Source de rayons X.**

Deux types de sources de rayons X sont utilisés en cristallographie des macromolécules biologiques : les sources classiques et le rayonnement synchrotron. Les sources classiques sont constituées d'anodes en métal pur bombardées par un faisceau d'électrons accélérés. Le spectre d'émission résultant présente des raies caractéristiques qui peuvent alors être utilisées pour le phénomène de diffraction. L'intensité du rayonnement reste cependant limitée, de même que le choix de la longueur d'onde de travail. Ce type de source est en général utilisé pour des recherches préliminaires en laboratoire.

L'utilisation du rayonnement synchrotron constitue l'une des révolutions majeures pour l'étude structurale des molécules biologiques. Les synchrotrons sont des anneaux dans lesquels des électrons (ou positrons) relativistes circulent en continu et sont accélérés à chaque passage. Ces électrons émettent alors un rayonnement blanc très intense et fortement polarisé dans le plan de l'anneau. Avec des monochromateurs, il est possible d'isoler la longueur d'onde désirée à partir d'une large gamme d'énergies. Le rayonnement synchrotron a permis de réaliser des progrès considérables dans le domaine de la cristallographie biologique : les temps d'exposition ont été réduits et l'augmentation de l'intensité du rayonnement a permis d'utiliser des cristaux de plus petite taille ; la résolution spatiale des taches de diffraction a été améliorée ; le rapport signal sur bruit a également augmenté, ce qui permet une plus haute résolution. Enfin, l'utilisation d'un rayonnement blanc permet de choisir la longueur d'onde adaptée à chaque étude. Les études cristallographiques décrites dans ce manuscrit ont été réalisées à partir des sources synchrotrons de l'European Synchrotron Radiation Facilities (ESRF) à Grenoble.

## **2. Préparation des échantillons.**

L'utilisation d'un rayonnement de type synchrotron très intense entraîne cependant une dégradation rapide des cristaux à température ambiante. C'est pourquoi, parallèlement à l'utilisation de ce rayonnement, une nouvelle technique a été développée qui consiste à enregistrer les données de diffraction à très basse température : la cryocrystallographie. Ce procédé consiste à mesurer la diffraction d'un cristal refroidi entre 100 et 140 Kelvins (K). Il a pour principal intérêt d'augmenter considérablement la durée de vie du cristal car à cette température, la progression des radicaux libres provoqués par les photons X dans le cristal est bloquée, limitant ainsi les dommages causés par l'irradiation au cours de l'acquisition des données.

Le refroidissement du cristal sans l'endommager reste le problème majeur de cette technique. En effet, la cryocrystallographie implique la congélation immédiate des cristaux à environ 140 K dans un flux d'azote gazeux ou d'éthane liquide. Les cristaux de protéine contenant une grande proportion de solvant (20 à 90 %), il faut éviter la formation de glace cristalline qui détruirait le cristal. Pour cela, le cristal est refroidi en présence d'une solution cryoprotectrice adaptée. Un bon cryoprotectant doit d'une part favoriser la formation de glace amorphe lorsqu'il est refroidi instantanément et d'autre part préserver le pouvoir diffractant du cristal. Les cryoprotectants les plus fréquemment utilisés sont le glycérol, l'éthylène glycol, les polyéthylènes glycols de faible poids moléculaire, en particulier le PEG 400, et le MPD. Les conditions de cryoprotection relatives aux conditions de cristallisation des solutions commerciales *Crystal Screen I* et *II* de Hampton Research sont désormais disponibles dans la littérature (McFerrin and Snell, 2002) et nous ont servi de base pour déterminer les cryoprotectants correspondants à d'autres conditions de cristallisation.

Les cristaux de Pa-aIF2 $\alpha_{D2-3}$  ont été cryoprotégés par trempage dans une solution 2.0 M sulfate d'ammonium, 25 % glycérol. Puis ils ont été congelés dans de l'éthane liquide et conservés ainsi jusqu'à l'enregistrement des données. L'isopropanol présent dans la solution de cristallisation a été omis dans la solution cryoprotectrice car il provoque l'agitation des cristaux dans la goutte qui rend leur prélèvement plus difficile. De plus, la présence d'isopropanol dans le cryoprotectant semble affecter la qualité des clichés de diffraction.

## **3. Acquisition et traitement des données pour la protéine native.**

Comme nous l'avons vu précédemment, la densité électronique d'un cristal est l'image, par transformée de Fourier, des facteurs de structures  $F_{hkl}$  qui caractérisent l'espace réciproque. Il est donc nécessaire de cartographier l'ensemble de cet espace réciproque pour

pouvoir reconstruire la densité électronique dans l'espace réel, ce qui revient à étudier les ondes diffractées par le cristal dans toutes les directions de l'espace. La méthode couramment utilisée en radiocristallographie est la méthode dite de rotation du cristal qui consiste à faire tourner le cristal d'un petit angle  $\Delta\phi$  autour d'un axe perpendiculaire au faisceau de rayons X. Le choix de cet angle d'oscillation est un compromis entre le nombre de clichés à enregistrer et la minimisation de la superposition des taches de diffraction. Un jeu de données complet est alors obtenu en balayant tout l'espace de diffraction. En général, du fait de la symétrie du cristal, il n'est pas nécessaire de balayer tout l'espace réciproque. Avant enregistrement du jeu de données, la détermination du groupe d'espace auquel appartient le cristal permet d'établir une stratégie d'exposition qui permet, en un minimum de clichés, d'obtenir le jeu de réflexions minimum qui permettra ensuite de retrouver l'ensemble des réflexions du cristal en appliquant les éléments de symétrie de son groupe d'espace.

Lorsqu'un jeu de données complet est enregistré, celui-ci est traité de manière à convertir les intensités brutes mesurées à partir des clichés de diffraction en intensité intégrées. Ce traitement se déroule en trois étapes :

- à partir d'un cliché de diffraction, détermination des paramètres de maille du cristal, de sa mosaïcité et de son orientation. Ces différents paramètres ainsi que ceux relatifs à l'acquisition des données (distance cristal-détecteur) sont ensuite affinés sur l'ensemble des clichés de diffraction.
- un fichier contenant toutes les réflexions prédites à partir des paramètres du cristal est alors créé. Les intensités sont ensuite intégrées par superposition des taches prédites sur les clichés de diffraction par ajustement de profil, ce qui permet d'améliorer la qualité de mesures surtout dans le cas où les réflexions sont de faible intensité.
- enfin, les différentes intensités sont mises à l'échelle. Pour cela, l'intensité de chaque réflexion indépendante est calculée à partir des mesures de toutes les réflexions équivalentes enregistrées sur l'ensemble des clichés. Des constantes d'échelle sont alors calculées entre chaque cliché de manière à homogénéiser les données.

Finalement, un programme donne la racine carrée des intensités, pour chaque réflexion utilisée, et permet d'obtenir un jeu de modules des facteurs de structure. Dans notre cas, il s'agit du programme *truncate* (Collaborative Computational Project n°4, 1994).

La qualité du jeu de données de diffraction est donnée par le facteur  $R_{\text{sym}}$  :

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

où  $I_{\text{hkl}}^i$  est la  $i$ ème mesure de la réflexion hkl et  $\langle I_{\text{hkl}} \rangle$  la valeur moyenne des N réflexions communes.

La collecte des jeux de données relatifs à la diffraction des rayons X par les cristaux d' $\alpha_{D2-3}$  a été réalisée à partir des sources de rayonnement synchrotron de l'ESRF à Grenoble. Un jeu de données à 2.26 Å a été collecté sur la ligne ID14-eh1, par la méthode de rotation du cristal. Les données de diffraction ont ensuite été traitées à l'aide des programmes *MOSFLM* (Leslie A.G.W., MRC Laboratory of Molecular Biology, Cambridge, Royaume-Uni) et mises à l'échelle avec *SCALA* (Collaborative Computational Project n°4, 1994). Un  $R_{sym}$  de 4.7 % a été obtenu pour ce jeu natif avec une valeur de 30.7 % dans la plus haute tranche de résolution (**Tableau 4.6**).

La résolution de cette structure a ensuite été tentée en utilisant la méthode du remplacement moléculaire. Les deux structures de  $\alpha_{D1-2}$  disponibles (Nonato *et al.*, 2002 ; Dhaliwal and Hoffman, 2003) ont été utilisées comme modèles dans la recherche d'une solution pour le domaine 2 d'aIF2 $\alpha_{D2-3}$ . Aucune solution de remplacement moléculaire n'a pu être trouvée par le programme *PHASER* (Storoni *et al.*, 2004). Au vu de ces résultats, une autre méthode de résolution a été envisagée à partir d'une protéine sélénométhionylée.

#### **4) Résolution d'une structure par la diffusion anormale multi-longueur d'onde (MAD).**

##### **1. Principe.**

Dans le cas général, le phénomène de diffraction des rayons X par les électrons des atomes du cristal correspond à une diffusion élastique, pour laquelle le rayonnement est diffusé à la même longueur d'onde que le rayonnement incident. Cependant le cristal peut contenir certains atomes pour lesquels l'énergie des rayons X incidents est proche de leur seuil d'absorption. Dans ce cas, il se produit un phénomène de résonance qui se traduit par un changement de longueur d'onde de l'onde diffusée, du fait de l'absorption d'une partie du rayonnement. Ce phénomène provoque la rupture de la loi de Friedel, c'est-à-dire que  $F_{hkl}$  et  $F_{(-h-k-l)}$  ne sont plus complexes conjugués (d'où  $I_{hkl} \neq I_{(-h-k-l)}$ ).

Dans ces conditions, le facteur de diffusion atomique  $f_j$  n'est plus un réel positif, mais un nombre complexe qui peut s'écrire :

$$f_j = f_j^{\circ} + f_j' + if_j''$$

où  $f_j^{\circ}$  est le facteur de diffusion "normal" de l'atome  $j$  et  $f_j'$  et  $f_j''$  constituent les composantes dites « anormales » du facteur de diffusion atomique, qui dépendent de la longueur d'onde incidente.

En l'absence de centrosymétrie des clichés de diffraction due à la rupture de la loi de Friedel, deux informations essentielles peuvent être directement mesurées à partir du jeu de données. La première information, appelée différence anormale ou différence de Bijvoet, est propre au diffuseur anormal introduit dans la molécule :

$$D_{\text{ano}} = |F_{\lambda i}(\mathbf{hkl})| - |F_{\lambda i}(-\mathbf{h}-\mathbf{k}-\mathbf{l})|$$

La seconde information correspond à la différence de norme entre les facteurs de structure moyens observés à chaque longueur d'onde :

$$\Delta F_{\Delta ij} = |F_{\lambda i} - F_{\lambda j}| \text{ avec } F_{\lambda i} = (|F_{\lambda i}(\mathbf{hkl})| + |F_{\lambda i}(-\mathbf{h}-\mathbf{k}-\mathbf{l})|)/2$$

La méthode MAD consiste à exploiter la différence observée entre les intensités diffractées à des longueurs d'onde différentes par un cristal possédant des diffuseurs anormaux. Cette différence étant relativement faible, les mesures expérimentales doivent être de très bonne qualité pour minimiser les erreurs lors de l'enregistrement des données. C'est pourquoi le rayonnement synchrotron est indispensable pour cette technique.

En théorie, deux jeux complets sont suffisants pour lever l'indétermination des phases (avec deux longueurs d'ondes, le calcul aboutit à un système de quatre équations à quatre inconnues). Dans la pratique, on effectue en général des mesures à 3 longueurs d'onde différentes pour optimiser le signal et augmenter la surdétermination du rapport du nombre d'observations sur le nombre de paramètres. Les courbes de variation de  $f_j'$  et de  $f_j''$  en fonction de la longueur d'onde sont déterminées juste avant la collecte des données MAD en enregistrant un spectre de fluorescence X de l'échantillon. En général, les deux principales longueurs d'onde choisies pour optimiser le signal correspondent au maximum de  $f_j''$  et au minimum de  $f_j'$ .

La détermination finale des phases nécessite de connaître la position exacte des diffuseurs anormaux au sein de la molécule. Cette étape fait appel à la fonction de Patterson :

$$P(\mathbf{uvw}) = \iiint_V \rho(\mathbf{xyz}) \times \rho(\mathbf{x+u, y+v, z+w}) d\mathbf{x}d\mathbf{y}d\mathbf{z}$$

qui est le produit de convolution de la densité électronique  $\rho(\mathbf{xyz})$  avec son symétrique par rapport à l'origine. Cette fonction est en réalité la transformée de Fourier de  $(F_{\mathbf{hkl}})^2$  et elle peut donc être directement calculée à partir des intensités mesurées sur les clichés de diffraction. Les maxima de cette fonction correspondent aux vecteurs interatomiques (intra et intermoléculaires). La comparaison des fonctions de Patterson observées pour le cristal natif

et pour le cristal dans lequel des diffuseurs anomaux ont été incorporés permet alors de déterminer la position de ces diffuseurs.

Dans la pratique, les erreurs de mesure sur les facteurs de structure nous obligent à déterminer une densité de probabilité  $P_j(\varphi_{hkl})$  pour chaque cristal dérivé  $j$  et pour chaque réflexion. L'utilisation de plusieurs cristaux dérivés permet alors de calculer une distribution totale de probabilité :

$$P(\varphi_{hkl}) = \prod_i P_i(\varphi_{hkl})$$

Le maximum de cette distribution correspond à la phase la plus probable et le centroïde de la distribution est appelé « phase la meilleure » ( $\varphi_{best}$ ). L'utilisation de cette phase dans le calcul des cartes de densité électronique donne l'erreur quadratique moyenne la plus faible. La qualité de l'information sur la phase est alors traduite par la figure de mérite  $m$  qui reflète l'étroitesse de la distribution de probabilité :

$$m = \frac{\int P(\varphi) e^{i\varphi} d\varphi}{\int P(\varphi) d\varphi}$$

La première carte de densité électronique qui est alors calculée est :

$$\rho(xyz) = (1/V) (\sum_h \sum_k \sum_l m_{hkl}) |F_{hkl}| e^{i\varphi_{best}} e^{-2i\pi(hx+ky+lz)}$$

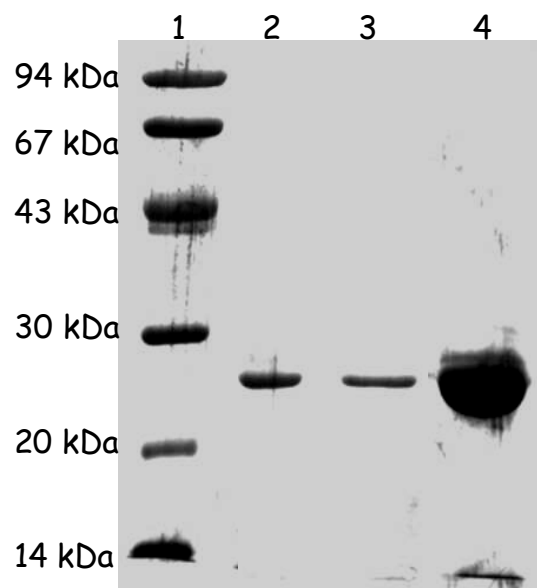
Aux longueurs d'onde utilisées en cristallographie des macromolécules biologiques, les atomes de faible numéro atomique présentent une diffusion anormale très faible qui est généralement négligée. Pour utiliser la méthode MAD comme technique de résolution de la structure de la protéine d'intérêt, il faut introduire des atomes plus lourds au sein des cristaux. Pour les protéines, c'est généralement le sélénium qui est utilisé car il peut être incorporé dans les polypeptides sous forme de sélénométhionine.

## **2. Obtention de la protéine $\alpha_{D2-3}$ sélénométhionylée, cristallisation et enregistrement des données.**

Nous avons choisi d'utiliser la méthode dite MAD (Multiwavelength Anomalous Diffraction) pour résoudre la structure de Pa-aIF2 $\alpha_{D2-3}$ . Cependant, la protéine  $\alpha_{D2-3}$  ne possède pas de méthionine dans sa séquence. Il a donc été nécessaire dans un premier temps de modifier la protéine pour qu'elle puisse incorporer des sélénométhionines. La forme tronquée de Pa-aIF2 $\alpha$  correspondant aux domaines 2 et 3 comporte 176 résidus.

L'incorporation de 2 méthionines, une dans chaque domaine, a donc été envisagée. Deux codons méthionine ont ainsi été insérés par mutagenèse dirigée au niveau des positions 106 (domaine 2) et 184 (domaine 3) de la protéine  $\alpha_{D2-3}$  (numérotation correspondant à celle de la sous-unité  $\alpha$  entière). Ces deux positions ont été choisies après avoir comparé les différentes séquences d' $\alpha$ IF2 $\alpha$  disponibles et identifié les positions pour lesquelles un résidu méthionine est présent (Voir **Annexe 3**). Ainsi, chez *M. jannaschii*, une méthionine occupe les positions correspondant à L106 et I184 chez *P. abyssi*.

Pour construire le double mutant  $\alpha_{D2-3}$ L106M-I184M, l'ADN codant pour  $\alpha_{D2-3}$  a été muté par PCR par la méthode *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene en utilisant deux amorces internes introduisant chacune l'une ou l'autre des deux mutations. Le plasmide doublement muté ainsi obtenu a été utilisé pour co-transformer la souche auxotrophe pour la méthionine B834(DE3) en présence du plasmide pSBETa. Une préculture de ces cellules transformées a été réalisée en milieu riche. Cette préculture a alors été utilisée pour ensemercer un litre de milieu minimum dans lequel la méthionine a été remplacé par de la sélénométhionine (voir la partie **Matériel et Méthodes** pour le protocole détaillé). La protéine sélénométhionylée a ensuite été purifiée selon le même protocole que la protéine native (**Figure 4.4**).



**Figure 4.4 :** Purification de la forme  $\alpha_{D2-3}$  sélénométhionylée de la sous-unité  $\alpha$ IF2 $\alpha$  de *P. abyssi*. Des aliquots de 20  $\mu$ L des différentes étapes de purification ont été déposés sur gel SDS-PAGE (12% en polyacrylamide). Piste 1 : marqueur de poids moléculaire ; Piste 2 : pic d'éluion de la première Q-Hiload ; Piste 3 : pic d'éluion de la seconde Q-Hiload; Piste 4 : témoin  $\alpha_{D2-3}$  natif.

Les cristaux d' $\alpha_{D2-3}$  sélénométhionylée ont été obtenus dans la même condition de cristallisation que les cristaux natifs. Leur apparition a cependant été plus lente puisque ces cristaux ont poussé en quatre semaines à 24°C. La cryoprotection a été réalisée avec la même

solution cryoprotectrice que pour la protéine native. Enfin, les données de diffraction ont été collectées sur la ligne ID23 de l'ESRF à Grenoble. Deux jeux de données à 2.8 Å ont pu être enregistrés à deux longueurs d'onde différentes correspondant respectivement au maximum et à l'inflexion du pic d'absorption du sélénium. Les conditions de collecte pour les cristaux natifs et sélénométhionylés sont reprises dans le **Tableau 4.5**.

	$\alpha_{D2-3}$ native		$\alpha_{D2-3}$ SeMet	
Source Synchrotron	ESRF	ESRF	ESRF	ESRF
Ligne	ID14-eh1	ID23	ID23	ID23
$\lambda$ (Å)	0.9340	0.9790	0.9756	0.9756
Détecteur	Quantum 4R CCD	MarMosaic 225 CCD	MarMosaic 225 CCD	MarMosaic 225 CCD
Angle d'oscillation	1°	1°	1°	1°
Distance au détecteur (mm)	200	250	250	250
Résolution max (Å)	2.14	2.6	2.6	2.6
Temps d'exposition (s)	30 (3 passages)	6 (3 passages)	6 (3 passages)	6 (3 passages)
Nombre de clichés	90	90	90	90

**Tableau 4.5 :** Conditions d'acquisition des données de diffraction des cristaux natifs et sélénométhionylés d'aIF2 $\alpha_{D2-3}$  de *P. abyssi*.

Comme pour le jeu natif, les données de diffraction relatives aux cristaux sélénométhionylés ont été traitées à l'aide des programmes *MOSFLM* (Leslie A.G.W., MRC Laboratory of Molecular Biology, Cambridge, Royaume-Uni) et *SCALA* (Collaborative Computational Project n°4, 1994). Les statistiques des jeux de données natifs et sélénométhionylés sont détaillées dans le **Tableau 4.6**.

### 3. Résolution de la structure.

La structure d'aIF2 $\alpha_{D2-3}$  a été résolue par la méthode MAD. Les deux jeux de données correspondant aux cristaux sélénométhionylés ont été utilisés comme paramètres d'entrée du programme *SOLVE* (Terwilliger and Berendzen, 1999). Deux sites correspondant aux deux atomes de sélénium attendus ont été identifiés à l'aide de ce programme qui procède par analyse des fonctions de Patterson. Les positions de ces deux atomes ainsi déterminées ont ensuite été utilisées pour déterminer les phases du jeu de réflexions correspondant à la protéine native. La solution fournie par *SOLVE* présente une figure de mérite  $m = 0.35$  à 2.8 Å. Une première carte de densité électronique a alors été tracée à partir de ce jeu initial de réflexions et de phases

Cristal	$\alpha$ D2-3 natif	$\alpha$ D2-3 sélénométhionylé	
Groupe d'espace	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2	
Paramètres de maille	a = 61.5 Å	a = 63.5 Å	
	b = 79.8 Å	b = 80.85 Å	
	c = 47.9 Å	c = 47.0 Å	
Résolution (Å)	2.26	2.8	2.8
Réflexions uniques	11532	5906	5896
Complétude (%)	99.7	94	93.6
Données anormales (%)	-	85.8	85.6
Redondance	3.4	3.7	3.6
R <sub>sym</sub> (I) (%) <sup>1</sup>	4.7 (30.7)	6.0 (36.3)	5.9 (40)
Figure de mérite	-	0.35 (de 20 à 2.8 Å)	

**Tableau 4.6 :** Statistiques du traitement des données. <sup>1</sup> Pour le R<sub>sym</sub>, les chiffres entre parenthèses correspondent aux statistiques dans la plus haute tranche de résolution.

### 5) Affinement du modèle.

L'affinement est la dernière étape du processus cristallographique et il consiste à optimiser le modèle issu de l'interprétation des cartes de densités électroniques par un processus mathématique qui minimise l'écart entre les modules des facteurs de structures observés et les modules des facteurs de structures calculés. De manière pratique, il s'agit d'un processus cyclique en deux étapes :

- une étape de reconstruction manuelle où l'on modifie le modèle en s'aidant des cartes de densités électroniques. Un nouveau jeu de facteurs de structure (modules et phases) est alors calculé à partir du nouveau modèle.
- une étape d'optimisation du modèle par calculs automatiques qui conduit à une minimisation de la différence entre  $F_{\text{calc}}$  et  $F_{\text{obs}}$ . Ces calculs aboutissent à un nouveau jeu de facteurs de structure qui est utilisé pour tracer une nouvelle carte de densité électronique.

Le déroulement de l'affinement cristallographique est contrôlé par l'évolution d'un facteur de désaccord cristallographique  $R_{\text{crist}}$  :

$$R_{\text{crist}} = \frac{\sum |F_{\text{hklobs}}| - k |F_{\text{hklcalc}}|}{\sum |F_{\text{hklobs}}|}$$

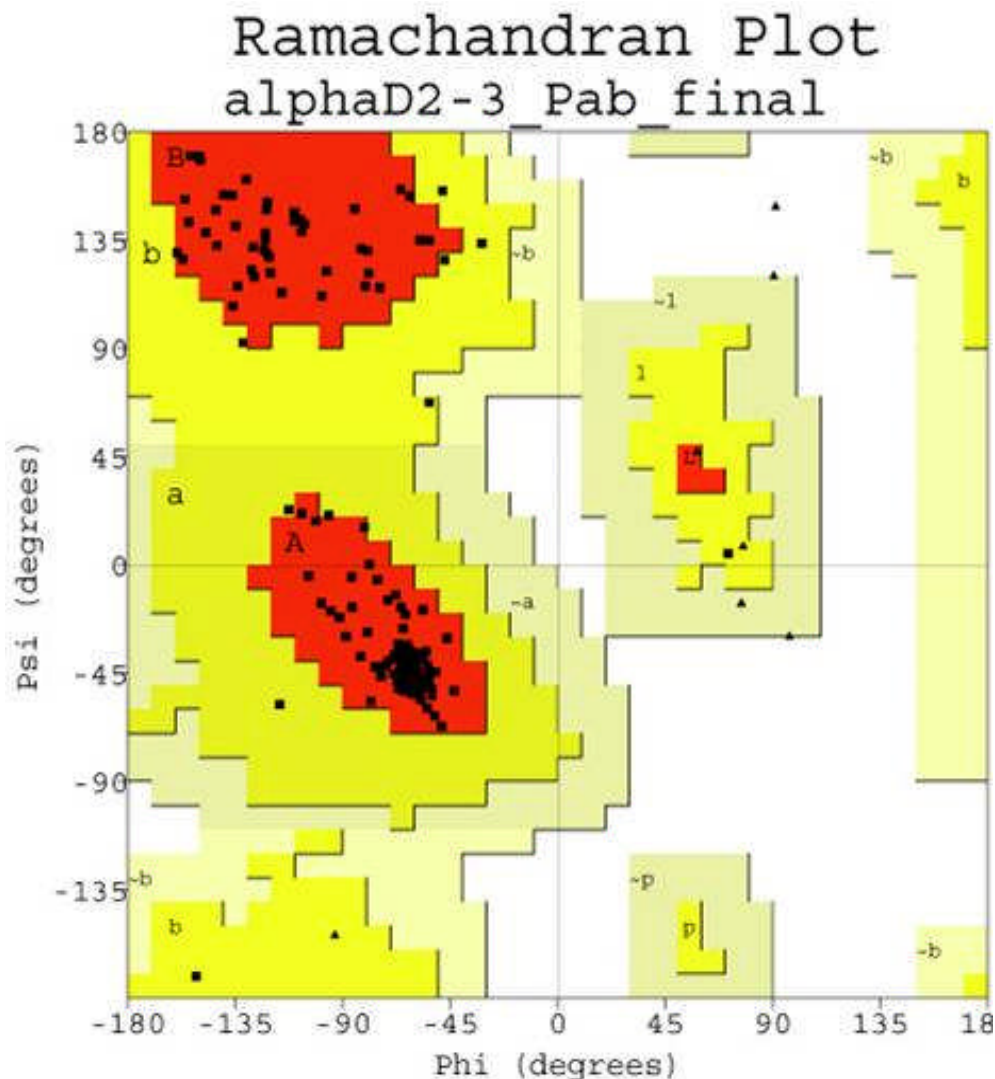
Cependant, ce facteur ne reflète pas l'exactitude du modèle car il ne prend pas en compte l'erreur sur les phases. Pour pallier ce défaut, la justesse du modèle est évaluée à l'aide d'un facteur de désaccord cristallographique particulier appelé  $R_{\text{libre}}$ . Ce facteur est calculé à partir de réflexions qui sont mises de côté dès le début du processus d'affinement et ne sont donc pas utilisées dans le processus d'optimisation (5 à 10% des réflexions en général).

Pour la résolution de la structure de Pa-aIF2 $\alpha_{\text{D2-3}}$ , la carte de densité électronique initiale a d'abord été améliorée par lissage de solvant à l'aide du programme *RESOLVE* (Terwilliger, 1999). La carte ainsi obtenue présentait une densité électronique suffisamment organisée pour positionner le modèle structural correspondant au domaine 2 de la protéine eIF2 $\alpha$  humaine (Nonato *et al.*, 2002). Ce polypeptide a ensuite été converti en une chaîne polyalanine à l'exception des résidus pour lesquels le même acide aminé est rencontré à la fois chez aIF2 $\alpha$  de *P. abyssi* et eIF2 $\alpha$  humaine. Le repositionnement des C $\alpha$  et des chaînes latérales des résidus du modèle a ensuite été effectué au fur et à mesure des cycles d'affinement. La construction du modèle a été effectuée à l'aide du programme *O* (Jones *et al.*, 1991) et l'affinement du modèle a été réalisé par différents cycles de minimisation et de dynamique moléculaire en utilisant le programme *CNS* (Brunger *et al.*, 1998).

<b>aIF2<math>\alpha_{\text{D2-3}}</math> de <i>P. abyssi</i></b>	
R / $R_{\text{libre}}$ (calculé sur 6% des réflexions)	23.4 / 28.1
Rmsd liaisons (Å) / Rmsd angles (°)	0.0077 / 1.32
Nombre de résidus	176
Nombre de molécules d'eau	65
<b>Facteurs B</b>	
Protéine (Å <sup>2</sup> )	45.9
Molécules d'eau (Å <sup>2</sup> )	51.4
<b>Diagramme de Ramachandran</b>	
Résidus dans la zone favorable (%)	95.6
Résidus dans les zones tolérées (%)	4.4
Résidus dans les zones généreusement tolérées (%)	0
Résidus dans les zones non admises (%)	0

**Tableau 4.7 :** Données cristallographiques globales du modèle établi pour aIF2 $\alpha_{\text{D2-3}}$  de *P. abyssi*. Ces statistiques proviennent de l'analyse du modèle final à l'aide du programme *PROCHECK* (Laskowski *et al.*, 1993).

La carte de densité ainsi affinée a très vite permis d'identifier les éléments de structure relatifs au domaine 3. Le domaine 3 a ainsi été reconstruit progressivement et l'ensemble du modèle a été affiné avec le programme *CNS*. Cette reconstruction a conduit à un modèle final qui présente un facteur R de 23.3% et un facteur  $R_{\text{libre}}$  de 28%. Le **Tableau 4.7** reprend les statistiques relatives au modèle construit.



**Figure 4.8 :** Diagramme de Ramachandran du modèle d' $\alpha F2\alpha_{D2-3}$  de *P. abyssi* affiné à 2.26 Å de résolution. Les zones en rouge représentent les zones favorables (95.6%), les zones en jaune vif indiquent les zones tolérées (4.4%) et les zones en jaune pâle correspondent aux régions généreusement tolérées. Enfin, les régions non admises sont indiquées en blanc. Les différents résidus sont symbolisés par des carrés à l'exception des glycines représentées par un triangle. Ce diagramme a été réalisé à l'aide du programme *PROCHECK* (Laskowski *et al.*, 1993).

La chaîne principale des domaines 2 et 3 de la protéine  $\alpha F2\alpha$  de *P. abyssi* a pu être construite quasi entièrement. La forme tronquée  $\alpha_{D2-3}$  regroupe les résidus 89 à 275 ainsi

qu'une méthionine en position 88. Le modèle cristallographique construit correspond quant à lui aux résidus 91 à 266. Les trois premiers résidus de la protéine ne sont pas visibles dans la densité électronique. D'autre part, la chaîne latérale du résidu 91 n'apparaît pas clairement dans la carte et ce résidu a par conséquent été modélisé par une alanine. L'extrémité C-terminale de la protéine (résidus 267 à 275) est également difficilement interprétable dans la carte de densité électronique. Cette région correspond à une extension très riche en lysines et l'absence d'une densité clairement interprétable à son endroit suggère la mobilité de cette zone au sein du cristal. Le reste de la protéine a pu être construit sans ambiguïté. Les chaînes latérales de tous les résidus ont été positionnées dans la densité électronique et tous les résidus adoptent une conformation favorable comme en témoigne le diagramme de Ramachandran (**Figure 4.8**). Enfin, le facteur d'agitation thermique moyen obtenu ( $B = 45.9 \text{ \AA}^2$  pour la protéine) s'accorde parfaitement avec la valeur déduite de la représentation de Wilson du jeu de données, qui est de  $45 \text{ \AA}^2$ . La représentation de Wilson est calculée par le programme *truncate* (Collaborative Computational Project n°4, 1994). Elle correspond au tracé du log des intensités en fonction de l'inverse du carré de la résolution. Cette courbe présente généralement une partie linéaire pour la zone correspondant aux résolutions en-dessous de 3-5  $\text{\AA}$  et le facteur d'agitation thermique est donné par la pente de cette partie linéaire.

## **6) Résolution de la structure d'aIF2 $\alpha$ de *P. abyssii*.**

### **1. Cristallisation d'aIF2 $\alpha$ et acquisition des données.**

La cristallisation de la sous-unité aIF2 $\alpha$  de *P. abyssii* a été réalisée manuellement à partir de la technique de la diffusion de vapeur en goutte suspendue en boîtes de 24 puits (gouttes de 4  $\mu\text{L}$  : 2  $\mu\text{L}$  de solution protéique + 2  $\mu\text{L}$  de réservoir ; réservoir : 500  $\mu\text{L}$ ). Une matrice d'échantillonnage réalisée au laboratoire a été utilisée pour la recherche des conditions de cristallisation de la protéine. Le sulfate d'ammonium a été employé comme agent précipitant. Les premiers essais de cristallisation ont été effectués à 24°C à partir d'une solution de protéine  $\alpha$  fraîchement purifiée suivant le protocole décrit et concentrée à 15  $\text{mg.mL}^{-1}$ . Au bout de 2 jours, des cristaux sont apparus dans les gouttes contenant 0.6 M sulfate d'ammonium, 0.83 M acétate de sodium pH 5.0.

La collecte des jeux de données relatifs à la diffraction des rayons X par ces cristaux a été réalisée à partir des sources de rayonnement synchrotron de l'ESRF à Grenoble (ligne ID14-eh1). Les cristaux ont été préalablement cryoprotégés par trempage dans une solution contenant 0.5 M sulfate d'ammonium, 0.83 M acétate de sodium pH 5.0, 40 % éthylène glycol (cryoprotectant), puis congelés dans de l'éthane liquide. Un jeu de données à 3.37  $\text{\AA}$  a ainsi été collecté sur la ligne ID14-eh1 de l'ESRF, par la méthode de rotation du cristal. Les

données de diffraction ont ensuite été traitées à l'aide des programmes *MOSFLM* (Leslie A.G.W., MRC Laboratory of Molecular Biology, Cambridge, Royaume-Uni) et *SCALA* (Collaborative Computational Project n°4, 1994). Les conditions d'acquisition pour les cristaux natifs sont reprises dans le **Tableau 4.9** et les statistiques du traitement de données dans le **Tableau 4.10**.

	$\alpha$ native
Source Synchrotron	ESRF
Ligne	ID14-eh1
$\lambda$ (Å)	0.9340
Détecteur	Quantum 4R CCD
Angle d'oscillation	1°
Distance au détecteur (mm)	260
Temps d'exposition (s)	15 (3 passages)
Nombre de clichés	80

**Tableau 4.9 :** Conditions d'acquisition des données de diffraction des cristaux natifs d'aIF2 $\alpha$  de *P. abyssi*.

Cristal	$\alpha$ natif
Groupe d'espace	P3 <sub>1</sub> 21
Paramètres de maille	a = 104.3 Å
	b = 104.3 Å
	c = 129.4 Å
Résolution (Å)	3.37
Réflexions uniques	11911
Complétude (%)	100
Redondance	5.4
Rsym(I) (%) <sup>1</sup>	5.8 (33.5)

**Tableau 4.10 :** Statistiques du traitement des données. <sup>1</sup> Pour le Rsym, les chiffres entre parenthèses correspondent aux statistiques dans la plus haute tranche de résolution.

## **2. Principe de la méthode du remplacement moléculaire.**

Le remplacement moléculaire est une méthode qui permet de résoudre la structure tridimensionnelle d'une molécule donnée (molécule A) lorsque l'on connaît la structure d'une molécule proche (molécule B), correspondant par exemple à un homologue de séquence.

Cette technique n'utilise que les modules des facteurs de structure et elle repose sur la comparaison des fonctions de Patterson de la molécule à étudier et de la molécule modèle. Ainsi, comme nous l'avons déjà signalé, la fonction de Patterson est en fait constituée des vecteurs interatomiques du cristal. Une partie de ces vecteurs correspond aux vecteurs intramoléculaires et peut généralement être séparée des vecteurs intermoléculaires sur un critère de longueur (les vecteurs intramoléculaires sont en général plus courts et donc plus proches de l'origine de la fonction de Patterson). L'hypothèse de base du remplacement moléculaire est que la molécule B modèle homologue structural de A peut être décrite par les mêmes vecteurs intramoléculaires que la molécule A, à l'exception près de leur position par rapport à l'origine et de leur orientation. Ainsi, si l'on donne à B une orientation et une position similaires à A au sein du cristal, les fonctions de Patterson qui décrivent les deux molécules seront analogues.

De manière pratique, le remplacement moléculaire consiste donc à trouver la matrice de rotation [R] et le vecteur de translation T qui permettent de « superposer » la fonction de Patterson calculée pour B à la fonction de Patterson déterminée expérimentalement pour A.

$$\mathbf{X}_B = [\mathbf{R}] \mathbf{X}_A + \mathbf{T}$$

Dans un premier temps, la matrice [R] est déterminée grâce à l'utilisation d'une fonction de rotation qui permet la recherche des maxima de recouvrement entre les deux Patterson. Le degré de similitude entre les deux fonctions de Patterson est alors évalué par l'intermédiaire d'un coefficient de corrélation :

$$\text{Corr}(\theta_1, \theta_2, \theta_3) = \int_{\mathbf{u}} \mathbf{P}_A(\mathbf{r}) \mathbf{P}_B([\mathbf{R}]\mathbf{r}) dV$$

où  $\mathbf{P}_A(\mathbf{r})$  représente la fonction de Patterson observée pour la molécule A en r

$\mathbf{P}_B([\mathbf{R}]\mathbf{r})$  représente la fonction de Patterson calculée pour la molécule B en [R]r, la matrice de rotation [R] étant construite sur les 3 angles  $\theta_1, \theta_2$  et  $\theta_3$ .

Les maxima de cette fonction de corrélation correspondent aux meilleures superpositions des deux fonctions. Les programmes de remplacement moléculaire classent alors les différentes matrices de rotation [R] correspondant à ces maxima en leur attribuant un « score », la meilleure solution présentant généralement un score élevé qui se détache nettement des solutions suivantes.

La fonction de translation est ensuite déterminée à partir des vecteurs intermoléculaires des fonctions de Patterson. La méthode consiste à rechercher les maxima de

corrélation entre la fonction de Patterson du cristal et la fonction de Patterson du modèle correctement orientée dans la maille, soit les maxima de la fonction suivante :

$$\mathbf{T}(\mathbf{t}) = \int_{\mathbf{v}} \mathbf{P}_A(\mathbf{u})\mathbf{P}_{B \text{ modifiée}}(\mathbf{u},\mathbf{t}) \, d\mathbf{u}$$

La fonction  $T(\mathbf{t})$  présente un maximum pour le vecteur  $\mathbf{t}$  correspondant à un vecteur intermoléculaire reliant deux molécules modèles correctement orientées dans la maille cristalline.

Une fois  $[R]$  et  $T$  déterminés et affinés, ils sont appliqués au jeu de coordonnées de la molécule A. Le nouveau jeu de coordonnées ainsi obtenu sert à calculer les phases qui sont ensuite associées aux facteurs de structure déterminés expérimentalement, pour calculer une première carte de densité.

### **3. Résolution de la structure.**

La structure d'aIF2 $\alpha$  de *P. abyssi* a été résolue par remplacement moléculaire à l'aide du programme *PHASER* (Storoni *et al.*, 2004) disponible sur l'interface graphique du CCP4 (Collaborative Computational Project N°4, 1994). Les cristaux de Pa-aIF2 $\alpha$  ont en fait été obtenus avant ceux de la forme tronquée Pa-aIF2 $\alpha_{D2-3}$ . Cependant, dans le cas de la protéine entière, aucun dérivé sélénométhionylé n'a pu être obtenu et les tentatives de phasage à l'aide de dérivés lourds se sont révélées infructueuses. Le remplacement moléculaire avait également été tenté sans succès à l'aide des structures modèles fournies par  $\alpha_{D1-2}$  humain et de levure (Nonato *et al.*, 2002 ; Dhaliwal and Hoffman, 2003). C'est en réalité la détermination de la structure du domaine 3 de Pa-aIF2 $\alpha$  qui a permis de pallier ce problème et d'obtenir enfin une solution correcte par remplacement moléculaire.

Deux structures ont été utilisées comme modèles pour tenter de phaser le jeu de données à 3.37 Å : il s'agit tout d'abord de la structure d'aIF2 $\alpha_{D2-3}$  de *P. abyssi* déterminée précédemment. Cette structure a servi de modèle pour les domaines 2 et 3 de la protéine  $\alpha$  entière. Dans un second temps, la structure d'eIF2 $\alpha$  de levure (Dhaliwal and Hoffman, 2003) a été utilisée comme modèle pour la recherche du domaine 1 de la protéine  $\alpha$ .

Dans un premier, le domaine 3, puis le domaine 2 d'aIF2 $\alpha$ , a été recherché en utilisant la structure d'aIF2 $\alpha_{D2-3}$ . Le programme *PHASER* attribue un score appelé *Zscore* aux différentes fonctions de rotation et de translation qu'il calcule pour superposer la Patterson modèle et la Patterson expérimentale. Une solution unique présentant un *Zscore* très détaché de 17.2 a ainsi été trouvée pour le domaine 3 d' $\alpha$ . Le domaine 2 a ensuite été identifié sous

forme d'une solution également unique avec un Zscore de 14.9. Le domaine 1 d'eIF2 $\alpha$  de levure a alors été utilisé pour rechercher le domaine 1 d'aIF2 $\alpha$  et une solution très détachée avec un Zscore de 10.9 a été à nouveau obtenue. Les solutions uniques relatives aux trois domaines d' $\alpha$  fournies par le programme *PHASER* ont alors permis de construire une première carte de densité électronique. Les trois solutions obtenues ont ensuite été tracées dans cette première carte. Les trois domaines ayant été recherchés séparément, les solutions obtenues n'appartiennent pas forcément à la même molécule au sein du cristal. La seconde étape a donc consisté à tracer les symétriques de ces différents domaines pour sélectionner les trois modèles appartenant à la même molécule et permettant la connexion entre les différents domaines d' $\alpha$ .

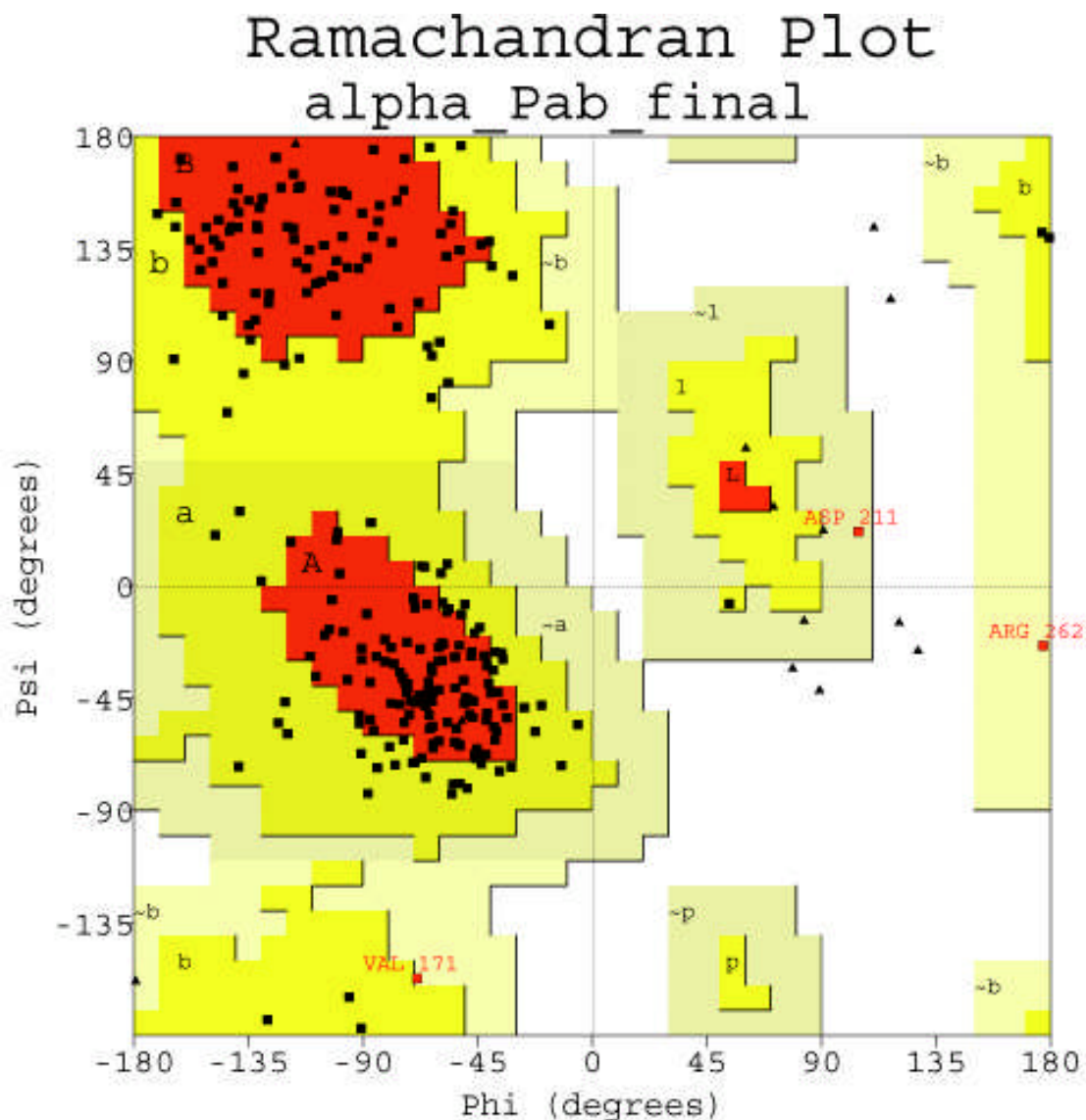
Dans le modèle ainsi obtenu, les résidus du domaine 1 correspondant à eIF2 $\alpha$  de levure ont été remplacés par des alanines à l'exception des résidus communs à aIF2 $\alpha$  de *P. abyssii* et à eIF2 $\alpha$  de levure. Le modèle a alors été progressivement reconstruit en alternant les cycles de repositionnement manuel à l'aide du programme *O* (Jones *et al.*, 1991) et les cycles d'affinement à l'aide du programme *CNS* (Brunger *et al.*, 1998). Pour les cycles de minimisation d'énergie, seules les données entre 12 et 3.37 Å (troncature des termes à basse résolution) et présentant un rapport F/ $\sigma$ (F) supérieur à 2 ont été prises en compte.

	aIF2 $\alpha$ de <i>P. abyssii</i>
R / R <sub>libre</sub> (calculé sur 6% des réflexions)	24.95 / 29.0
Rmsd liaisons (Å) / Rmsd angles (°)	0.0094 / 1.49
Nombre de résidus	261
<b>Facteurs B</b>	
Protéine (Å <sup>2</sup> )	105
<b>Diagramme de Ramachandran</b>	
Résidus dans la zone favorable (%)	69.9
Résidus dans les zones tolérées (%)	28.8
Résidus dans les zones généreusement tolérées (%)	1.3
Résidus dans les zones non admises (%)	0

**Tableau 4.11 :** Données cristallographiques globales du modèle établi pour aIF2 $\alpha$  de *P. abyssii*. Ces statistiques proviennent de l'analyse du modèle final à l'aide du programme *PROCHECK* (Laskowski *et al.*, 1993).

En fin d'affinement, le modèle a été amélioré en affinant les valeurs du facteur B par groupes d'atomes. Un facteur B moyen de 105 Å<sup>2</sup> a ainsi été obtenu. Une valeur relativement

élevée avait déjà été calculée à partir de la représentation de Wilson ( $B = 80 \text{ \AA}^2$ ). Cette forte valeur reflète en fait le faible pouvoir diffractant de ces cristaux et le fort pourcentage de solvant contenu dans la maille cristalline (76%). Cette reconstruction a finalement conduit à un modèle qui présente un facteur R de 23.8% et un facteur  $R_{\text{libre}}$  de 28.9%. Le **Tableau 4.11** reprend les statistiques relatives au modèle construit.



**Figure 4.12:** Diagramme de Ramachandran du modèle d'aIF2 $\alpha$  de *P. abyssi* affiné à 3.37 Å de résolution. Les zones en rouge représentent les zones favorables (69.9%), les zones en jaune vif indiquent les zones tolérées (28.8%) et les zones en jaune pâle correspondent aux régions généreusement tolérées (1.3%). Enfin, les régions non admises sont indiquées en blanc. Les différents résidus sont symbolisés par des carrés à l'exception des glycines représentées par un triangle. Ce diagramme a été réalisé à l'aide du programme PROCHECK (Laskowski *et al.*, 1993).

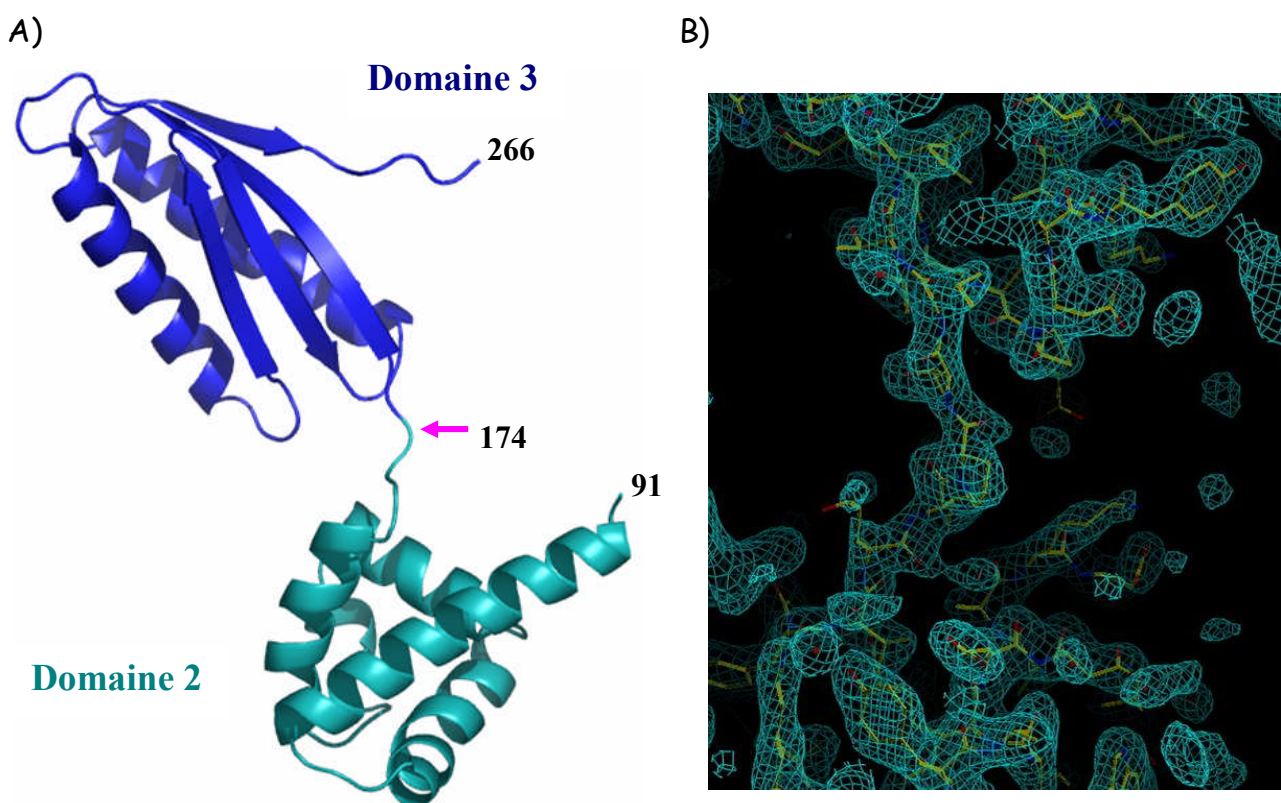
Le modèle cristallographique construit correspond aux résidus 5 à 265 d'aIF2 $\alpha$ . Les quatre premiers résidus de la protéine ne sont pas visibles dans la densité électronique. D'autre part, comme pour la forme tronquée  $\alpha_{D2-3}$ , les positions des résidus C-terminaux (résidus 266 à 275) n'ont pas pu être attribuées du fait d'une densité électronique non interprétable dans cette région. Le reste de la protéine a pu être construit sans ambiguïté et les chaînes latérales des différents résidus ont été positionnées dans la densité électronique. Enfin, le diagramme de Ramachandran (**Figure 4.12**) ne comporte que trois résidus pour lesquels le couple de valeurs pour les angles ( $\varphi, \psi$ ) se situe dans une zone généreusement tolérée : il s'agit des résidus Val 171, Asp 211 et Arg 262.

## **Partie B : Structure de la sous-unité aIF2 $\alpha$ de l'archée *Pyrococcus abyssi*.**

### **1) Description générale.**

#### **1. Structure de Pa-aIF2 $\alpha$ <sub>D2-3</sub>.**

La structure d'aIF2 $\alpha$ <sub>D2-3</sub> ainsi déterminée à 2.26 Å de résolution regroupe les résidus 91 à 266 de la protéine  $\alpha$  entière. Cette structure comprend deux domaines nettement séparés, reliés par un polypeptide en conformation étendue et parfaitement défini dans la densité électronique (**Figure 4.13**).



**Figure 4.13 :** Structure de la forme tronquée aIF2 $\alpha$ <sub>D2-3</sub> du facteur 2 de démarrage de la traduction de l'archée *P. abyssi* à 2.26 Å de résolution. A) Figure dessinée avec *PyMOL*. B) Vue de la densité électronique autour des résidus qui connectent les domaines 2 et 3 de la sous-unité  $\alpha$  réalisée avec *O* (Jones *et al.*, 1991); le polypeptide connecteur est parfaitement défini.

Le premier domaine correspond au domaine 2 (résidus 91 à 171) de la protéine aIF2 $\alpha$ . Il est composé de 5 hélices  $\alpha$  antiparallèles très proches les unes des autres formant un assemblage compact. Ce domaine est parfaitement superposable au domaine 2 des protéines

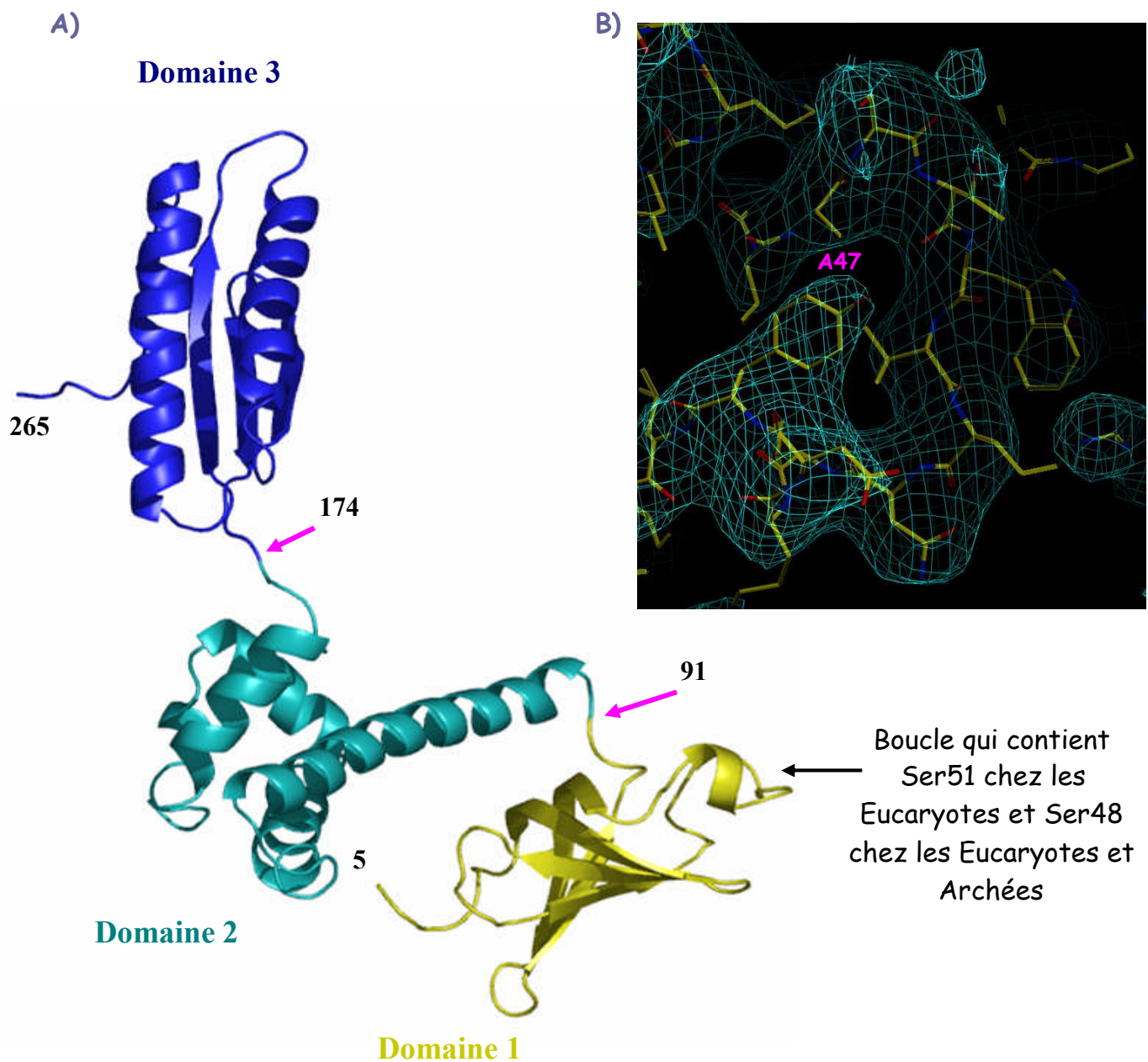
eIF2 $\alpha$  humaine et de levure (Nonato *et al.*, 2002 ; Dhaliwal and Hoffman, 2003). Un écart quadratique moyen (rmsd) de 1.6 Å pour la comparaison des positions de 78 carbones C $\alpha$  est obtenu lors de la superposition de ce domaine avec l'un ou l'autre des domaines 2 de la protéine humaine ou de levure. Le domaine C-terminal (domaine 3, résidus 176 à 275) est replié selon une topologie de type  $\beta\alpha\beta\beta\alpha\beta$ . Les deux hélices  $\alpha$  sont situées du même côté par rapport au feuillet  $\beta$  formé par les brins  $\beta$ 7,  $\beta$ 8,  $\beta$ 6 et  $\beta$ 9. En conséquence, la structure de ce domaine 3 d'aIF2 $\alpha$  présente un pseudo axe de symétrie interne d'ordre 2. Les deux molécules engendrées par cet axe de symétrie peuvent être superposées avec un rmsd de 2.0 Å pour 80 C $\alpha$  comparés.

## **2. Structure de la protéine intacte Pa-aIF2 $\alpha$ .**

La structure d'aIF2 $\alpha$  déterminée à 3.37 Å de résolution regroupe les résidus 5 à 265 de la protéine. Malgré une résolution assez faible, les 3 domaines structuraux ainsi que la région reliant les domaines 2 et 3 d' $\alpha$  sont parfaitement identifiables dans la densité électronique (**Figure 4.14**). L'arrangement en motif de type OB-fold du domaine 1 des protéines  $\alpha$  eucaryotiques est retrouvé dans cette structure. Les domaines 2 et 3 présentent quant à eux la même structure que pour la forme tronquée aIF2 $\alpha_{D2-3}$ .

La structure du module correspondant aux domaines 1 et 2 de la protéine a été superposée aux deux structures d'eIF2 $\alpha_{D1-2}$  disponibles dans la littérature. Un rmsd de 1.7 Å a été obtenu pour 152 C $\alpha$  comparés avec les domaines correspondant d'eIF2 $\alpha$  de levure (Dhaliwal and Hoffman, 2003) et un rmsd de 1.8 Å a été mesuré pour 143 C $\alpha$  comparés avec les domaines correspondant d'eIF2 $\alpha$  humaine (Nonato *et al.*, 2002). Ainsi, la position relative des domaines 1 et 2 d' $\alpha$  est conservée chez les Eucaryotes et chez les Archées (**Figure 4.15**).

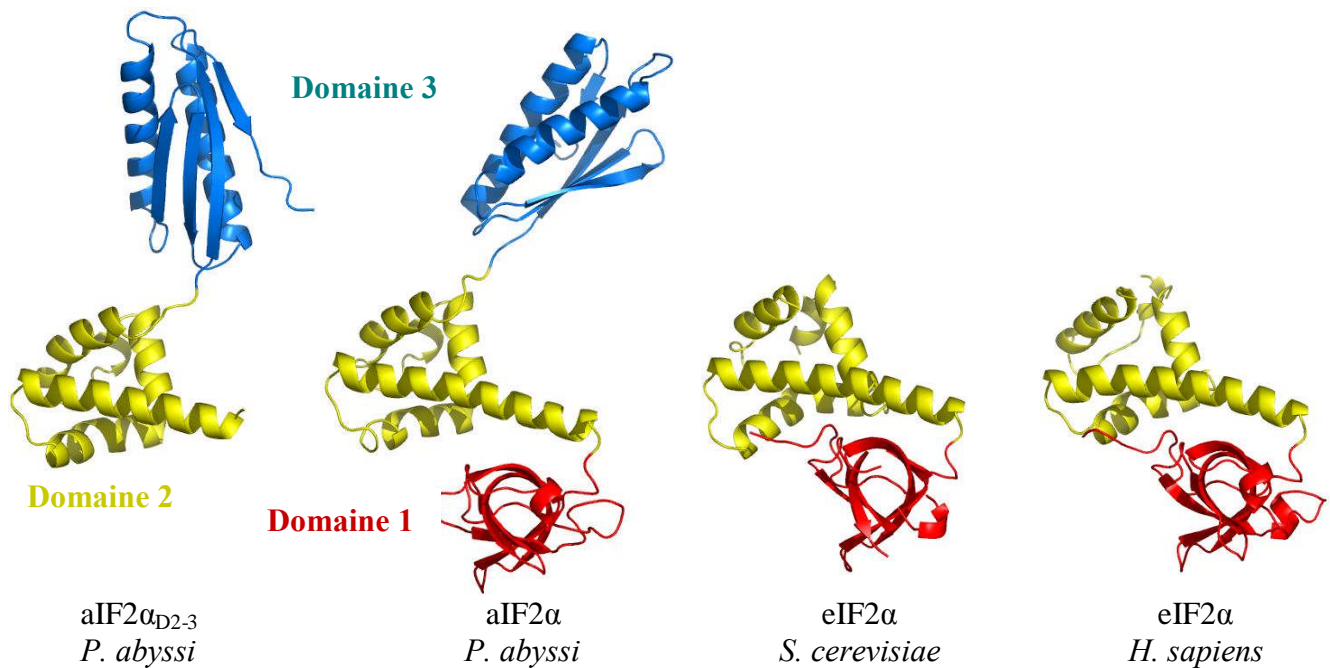
La boucle du domaine 1 contenant les petites hélices  $\alpha_A$  et  $\alpha_B$  correspond à la boucle qui porte la sérine 51 chez les Eucaryotes. Chez les Archées, aucun homologue de la sous-unité catalytique d'eIF2B n'a pu être mis en évidence jusqu'à présent et la sérine correspondant à la position Ser51 n'est pas conservée (alanine chez Pa-aIF2 $\alpha$ ) (Voir **Figure 4.19** et **Annexe 3**). Cependant, la boucle connectant les brins  $\beta$ 3 et  $\beta$ 4 dans la structure d'aIF2 $\alpha$  de *P. abyssi* adopte une conformation identique à celle observée dans la structure d'eIF2 $\alpha$  de levure. De plus, cette boucle contient également un résidu sérine (Ser48) très conservé chez les Archées (**Figure 4.19**) qui, chez *P. horikoshii*, est phosphorylé *in vitro* par l'homologue archéen de PKR (Tahara *et al.*, 2004).



**Figure 4.14 :** Structure de la sous-unité  $\alpha$  entière du facteur 2 de démarrage de la traduction de l'archée *P. abyssi* à 3.37 Å de résolution. A) Figure dessinée avec *PyMOL*. B) Vue de la densité électronique autour de la boucle du domaine 1 qui, chez les Eucaryotes, porte la sérine 51 phosphorylable. Vue réalisée avec *O* (Jones *et al.*, 1991) La position équivalente chez Pa- $\alpha$ IF2 $\alpha$  (Ala47) est indiquée.

### 3. Orientation relative des 3 domaines de la sous-unité $\alpha$ .

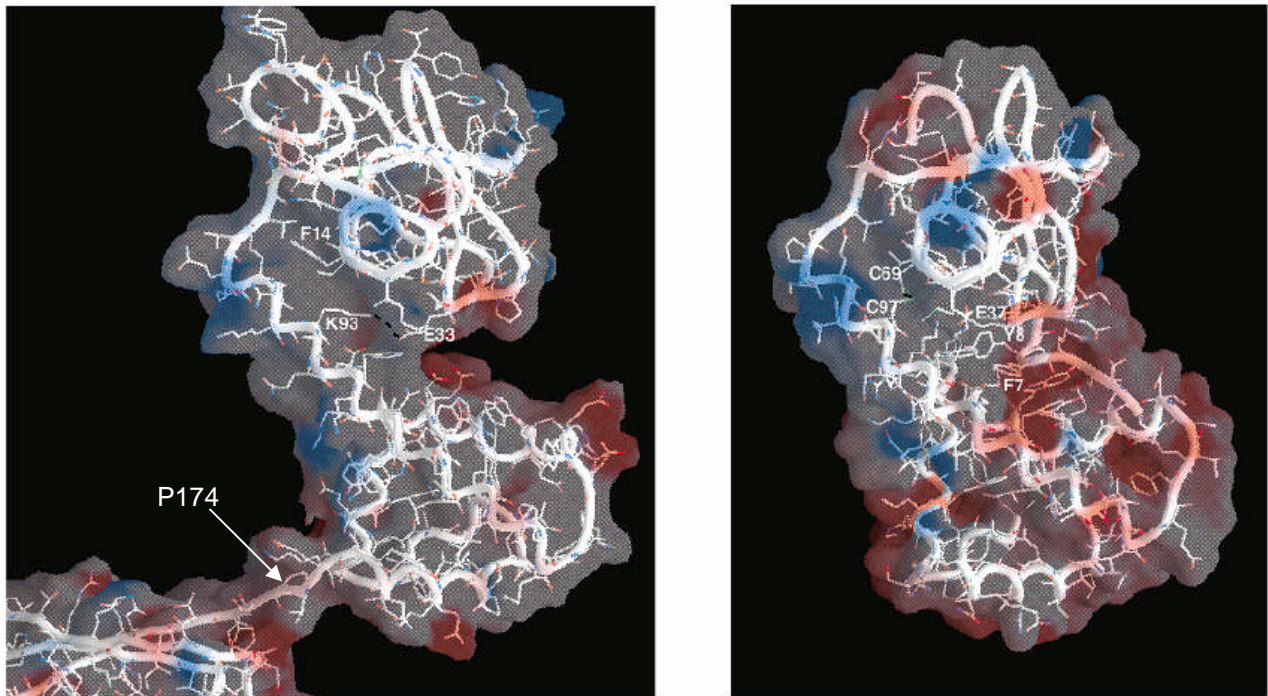
Comme l'a révélé la comparaison des structures d' $\alpha$ IF2 $\alpha$  avec les structures des eIF2 $\alpha$  humaine et de levure, les domaines 1 et 2 de ces protéines adoptent une orientation relative qui est conservée chez les Eucaryotes et chez les Archées. La **Figure 4.15** illustre cette conservation.



**Figure 4.15 : Organisation structurale des différents domaines des protéines e- $\alpha$ IF2 $\alpha$ .** Les structures de la forme tronquée  $\alpha_{D2-3}$  d' $\alpha$ IF2 $\alpha$  de *P. abyssi*, de la protéine entière  $\alpha$ IF2 $\alpha$  de *P. abyssi*, des domaines 1 et 2 d' $\alpha$ IF2 $\alpha$  de *S. cerevisiae* (Dhaliwal and Hoffman, 2003) et des domaines 1 et 2 d' $\alpha$ IF2 $\alpha$  humaine (Nonato *et al.*, 2002 ; Ito *et al.*, 2004) sont représentées en fonction de leur structure secondaire. Les trois domaines structuraux d' $\alpha$  sont colorés comme suit : le domaine 1 (résidus 1 à 84 chez *P. abyssi*) est dessiné en rouge, le domaine 2 (résidus 85 à 171 chez *P. abyssi*) est en jaune et le domaine 3 (résidus 172 à 266 chez *P. abyssi*) est représenté en cyan. Les domaines 2 de chaque structure ont été préalablement superposés dans le programme *O* (Jones *et al.*, 1991). Puis, chaque protéine a été dessinée en gardant la même orientation. Cette figure montre clairement que les domaines 1 et 2 adoptent une orientation respective constante parmi ces différentes structures tandis que l'orientation du domaine 3 varie. Cette figure a été construite à l'aide du programme Setor (Evans, 1993).

Plusieurs interactions permettent de maintenir les deux domaines dans des orientations relatives fixes. Ainsi, dans la structure des deux protéines  $\alpha$ IF2 $\alpha$  eucaryotiques, l'espace séparant les domaines 1 et 2 est occupé par une extension N-terminale de la protéine qui contient deux résidus (Phe7 et Tyr8 dans la numérotation de la levure) hautement conservés dans les différentes séquences d' $\alpha$ IF2 $\alpha$  connues à ce jour (Nonato *et al.*, 2002 ; Dhaliwal and Hoffman, 2003 ; Ito *et al.*, 2004). Les chaînes latérales de ces deux résidus aromatiques assurent des interactions de type Van der Waals qui permettent de ponter les deux domaines N-terminaux de la protéine. De plus, le résidu Tyr8 interagit avec le résidu Glu37 (strictement conservé dans toutes les séquences eucaryotiques d' $\alpha$ ) et avec la Ser104 (conservée dans 34 séquences sur 42), ce qui renforce la stabilité de l'interface et le maintien des deux domaines dans une position fixe l'un par rapport à l'autre (**Figure 4.16**). D'autre part, chez les animaux,  $\alpha$ IF2 $\alpha$  contient deux cystéines parfaitement conservées (aux positions C69 et C97 selon la numérotation de la protéine humaine). Ces deux cystéines forment un pont dissulfure qui relie ainsi les deux domaines (Nonato *et al.*, 2002). Chez les autres Eucaryotes, seule la seconde cystéine est conservée. Ainsi, dans la structure d' $\alpha$ IF2 $\alpha$  de levure, cette cystéine pointe vers

une valine qui vient ainsi se substituer à l'autre cystéine rencontrée chez les animaux (Dhaliwal and Hoffman, 2003).

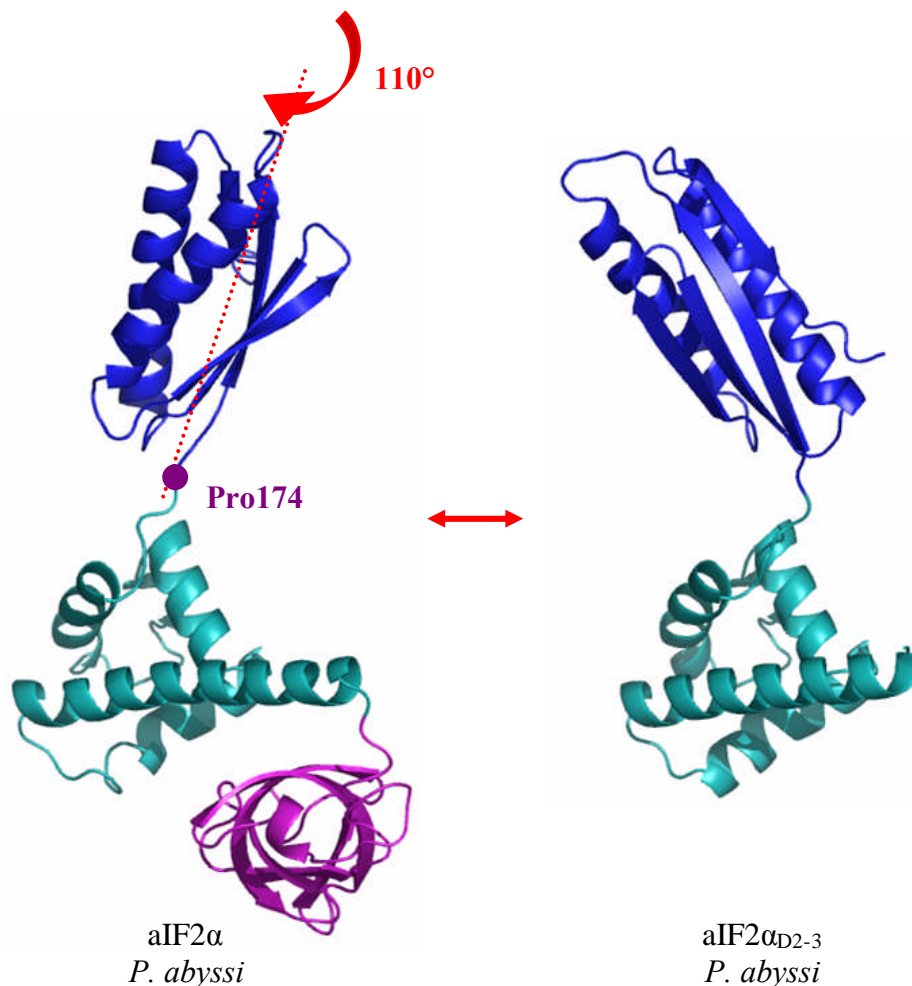


**Figure 4.16 :** Cartographie de la surface d'aIF2 $\alpha$  de *P. abyssi* (image de gauche) et de la surface du fragment N-terminal d'eIF2 $\alpha$  humaine (Nonato *et al.*, 2002) (image de droite). Ces figures représentent les potentiels électrostatiques de la surface de ces protéines. Ils ont été calculés avec le programme *Delphi* (Nicholls and Honig, 1991) et tracés avec le programme *Grasp* (Nicholls *et al.*, 1991). Les régions chargées négativement sont représentées en rouge et les zones chargées positivement sont dessinées en bleu. Les résidus indiqués correspondent aux résidus impliqués dans l'interface entre les domaines 1 et 2 des protéines e-aIF2 $\alpha$ .

Dans la structure d'aIF2 $\alpha$ , l'extension N-terminale rencontrée chez les Eucaryotes est absente et l'espace entre les deux domaines est de ce fait plus large. Cependant, l'interface est stabilisée par un pont salin qui s'établit entre les résidus Lys93 de l'hélice  $\alpha$ 4 et Glu33 dans la boucle reliant  $\beta$ 3 et  $\beta$ 4 (**Figure 4.16**). Ces deux résidus sont hautement conservés dans toutes les séquences d'aIF2 $\alpha$  d'Archées. La seule exception est celle de la protéine  $\alpha$  de l'archée *Halobacterium salinarium*. Cependant, des résidus de fonctionnalité proche sont trouvés aux positions correspondantes (**Figure 4.19 et Annexe 3**). Ainsi, une glutamine vient remplacer le résidu Glu33 et une thréonine occupe la position du résidu Lys93. Une interaction de même nature semble donc probable chez cet organisme. Chez aIF2 $\alpha$  de *P. abyssi*, le groupe amine du résidu Lys93 interagit également avec le groupe carbonyle de la chaîne principale au niveau du résidu Phe14. Ainsi, une même orientation est assurée entre les domaines 1 et 2 de la sous-unité  $\alpha$  chez les Eucaryotes et chez les Archées, bien que le jeu d'interactions mis en

œuvre dans chaque cas soit totalement différent. De plus, le caractère acide de l'interface est conservé (**Figure 4.16**).

L'orientation du domaine 3 est quant à elle tout autre. En effet, si l'on superpose le domaine 2 de la structure d'aIF2 $\alpha$  entière et celui de la structure de la forme tronquée  $\alpha_{D2-3}$ , l'orientation relative du domaine 3 par rapport au domaine 2 diffère alors considérablement d'une structure à l'autre (**Figure 4.17**).



**Figure 4.17.** Comparaison de l'orientation relative du domaine 3 d'aIF2 $\alpha$  par rapport aux domaines 1 et 2 pour les deux structures d'aIF2 $\alpha$  et aIF2 $\alpha_{D2-3}$  de *P. abyssi*. Les domaines 2 de chacune de ces deux structures ont été superposés à l'aide du programme *O* (Jones *et al.*, 1991). Les structures ont ensuite été dessinées à l'aide du logiciel *PyMOL* en gardant la même orientation.

La **Figure 4.17** révèle une rotation du domaine 3 d'aIF2 $\alpha$  de près de 110° entre les deux structures. De plus, aucun contact n'est assuré entre les domaines 2 et 3 de la protéine, aussi bien au sein de la protéine entière que dans le fragment  $\alpha_{D2-3}$ . Cette absence d'interaction entre les deux domaines C-terminaux d' $\alpha$  laisse donc libre champ au mouvement de rotation du domaine 3.

Chez  $eIF2\alpha$ , le mouvement entre les domaines 2 et 3 est articulé autour du résidu Pro174 (**Figures 4.16 et 4.17**) et crée une large variation de l'angle dièdre  $N-C\alpha_{173}-C-N_{174}$ . Ce résidu proline est très fréquemment rencontré à la position 173 ou 174 dans les séquences d' $eIF2\alpha$ . En effet, 56 séquences sur les 66 séquences comparées possèdent une telle proline (**Figure 4.19**). Ainsi, la mobilité du domaine 3 d' $\alpha$  par rapport au bloc constitué par les domaines 1 et 2 de la protéine semble être une caractéristique importante du facteur  $eIF2$ , à la fois chez les Eucaryotes et chez les Archées.

## **2) Homologies structurales et de séquence.**

### **1. Homologies structurales.**

#### **• Domaine 1 d' $eIF2\alpha$ .**

Le domaine 1 d' $eIF2\alpha$  présente un motif de type OB-fold. Cette topologie se retrouve dans de nombreuses protéines possédant des capacités de liaison des acides nucléiques comme IF1, la protéine ribosomale S1, diverses polynucléotides phosphorylases ou aminoacyl-ARNt synthétases (Gribskov, 1992). Dans tous les cas, le site de fixation des ARNs est formé par la partie concave du tonneau  $\beta$  (**Figure 4.18**).

De manière générale, la boucle qui porte la Ser51 phosphorylable chez les Eucaryotes vient prolonger le motif de liaison des ARNs du domaine 1 d' $eIF2\alpha$ . Une relation entre les propriétés de liaison des ARNs du domaine N-terminal d' $eIF2\alpha$  et la régulation de l'échange GDP – GTP au sein d' $eIF2\gamma$  pourrait donc être envisagée.

#### **• Domaine 3 d' $eIF2\alpha$ .**

Pour le domaine 3 d' $eIF2\alpha$ , une recherche d'homologues structuraux a été entreprise en utilisant la base de données DALI ([www.ebi.ac.uk/dali](http://www.ebi.ac.uk/dali)). Plus de 700 protéines ont été identifiées comme présentant une homologie structurale avec le domaine C-terminal d' $eIF2\alpha$  avec un Zscore supérieur à 2.0. Ce module structural qui adopte une topologie de type  $\beta\alpha\beta\beta\alpha\beta$  semble donc très répandu. Il se retrouve ainsi dans un large panel de protéines englobant à la fois des protéines impliquées dans le processus de traduction mais également des protéines ayant des fonctions biologiques sans aucun rapport avec la traduction. Ce module correspondant au domaine 3 d' $\alpha$  peut constituer un domaine particulier d'une protéine mais il peut également exister en tant qu'entité isolée.

Du fait de sa très large représentation au sein des protéines impliquées dans la traduction, ce motif a été désigné comme un motif fondamental de liaison des ARNs qui aurait émergé très tôt du monde ARN originel (Cousineau *et al.*, 1997 ; Jue *et al.*, 1980). Ce motif se retrouve dans 2 protéines liant les ARNs.

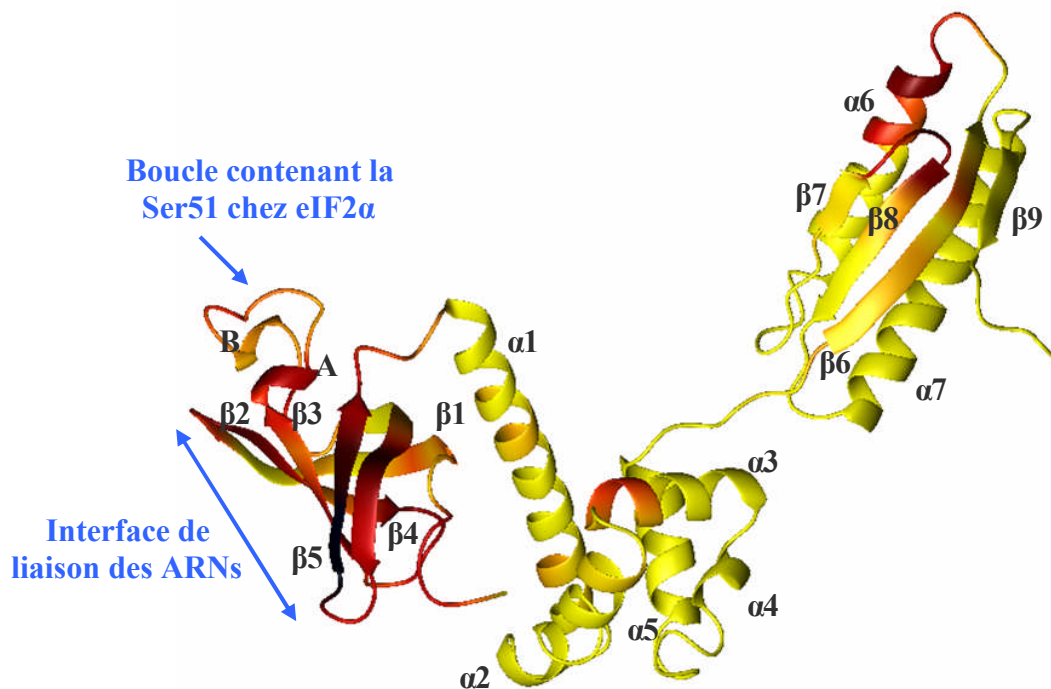
Parmi les protéines qui présentent le plus fort taux d'homologie avec  $\alpha_{D3}$ , on trouve :

- le domaine 3 de la translocase eEF2 (code d'entrée à la PDB : 1n0v ; Zscore = 7.2 ; rmsd = 1.9 Å pour 63 atomes comparés ; Jorgensen *et al.*, 2003),
- la protéine ribosomale S6 (code d'entrée à la PDB : 1ris ; Zscore = 6.4 ; rmsd = 2.1 Å pour 69 atomes comparés ; Lindhal *et al.*, 1994),
- le domaine C-terminal de la sous-unité  $\alpha$  du facteur d'échange eEF1B (code d'entrée à la PDB : 1b64 ; Zscore = 6.5 ; rmsd = 2.1 Å pour 72 atomes comparés ; Perez *et al.*, 1999).

## **2. Homologies de séquence.**

Pour mettre en évidence les régions fonctionnellement déterminantes des protéines e/aIF2 $\alpha$ , nous avons entrepris une recherche de conservation de séquences parmi toutes les séquences connues d'e/aIF2 $\alpha$  eucaryotiques ou archéennes. 66 séquences ont ainsi été isolées. Les séquences polypeptidiques ont d'abord été alignées à l'aide du programme *CLUSTAL X* (Thompson *et al.*, 1997). La **Figure 4.19** représente cet alignement. Seules quelques séquences caractéristiques ont été indiquées, cependant les colorations en fonction de la conservation tiennent compte de l'ensemble des 66 séquences (voir la légende de la figure pour plus de détails). Dans un second temps, la conservation de chaque position a été étudiée à l'aide du programme *Plotsimilarity* du groupe de programmes GCG disponible sur le site d'Infobiogen ([www.infobiogen.fr](http://www.infobiogen.fr)). La protéine aIF2 $\alpha$  a alors été redessinée et colorée en fonction de la conservation de chaque résidu, un dégradé allant du bleu au jaune permettant de passer des zones de très forte conservation aux zones où la nature des résidus est divergente d'une protéine à l'autre (**Figure 4.18**).

Sur cette figure, deux régions présentant les meilleurs taux de conservation se détachent très nettement. La première zone se situe dans le domaine 1 d' $\alpha$  et est centrée sur le site canonique de liaison des ARNs. La seconde zone se situe à l'autre extrémité de la protéine, dans le domaine 3 d' $\alpha$ . Elle est formée par la partie C-terminale du brin  $\beta_6$ , le début de l'hélice  $\alpha_6$ , la boucle connectant les brins  $\beta_7$  et  $\beta_8$  et le début du brin  $\beta_8$ . Le domaine 2 présente une plus grande variabilité et aucune zone particulièrement conservée n'a pu être mise en évidence dans le domaine central de la protéine. Seule la partie N-terminale de l'hélice  $\alpha_3$  présente un taux de conservation légèrement supérieur à la moyenne. Cette zone du domaine 3 pourrait être celle qui est en interaction avec la sous-unité aIF2 $\gamma$ .



**Figure 4.18 :** Positions des résidus conservés sur la structure d'eIF2 $\alpha$  de *P. abyssi*. La conservation de 66 séquences d'eIF2 $\alpha$  a été analysée comme décrit dans la partie Matériels et Méthodes. Les résidus présentant un taux de conservation inférieur à la moyenne sont tracés en jaune. Les résidus ayant un taux de conservation allant d'un niveau moyen au niveau maximal sont colorés suivant un dégradé partant du jaune au bleu de manière à ce que la couleur rouge représente une conservation intermédiaire entre le niveau moyen et le niveau maximal. Cette figure a été dessinée à l'aide du programme *Molmol* (Koradi et al., 1996).

### **3. Implications fonctionnelles.**

Les résultats de ces comparaisons, ajoutés aux données biochimiques obtenues dans le **Chapitre III**, indiquent qu'eIF2 $\alpha$  contient deux domaines de liaison des ARNs mobiles l'un par rapport à l'autre et jouant des rôles distincts.

Ainsi, le domaine 3 d' $\alpha$  est responsable de la liaison à la sous-unité  $\gamma$  et sa seule présence permet de doter la sous-unité  $\gamma$  d'une affinité optimale pour le Met-ARNt<sub>i</sub><sup>Met</sup>. Les expériences de retard de migration sur gel montrent que ce domaine possède des propriétés générales de liaison des acides nucléiques qui suggèrent qu'il pourrait contribuer à un renforcement de l'affinité pour l'ARNt par une interaction directe avec le polynucléotide. D'autre part, les expériences de retard de migration sur gel montrent que le domaine 1 présente également des propriétés générales de fixation des ARNs. Le motif OB-fold n'interagit pas avec la sous-unité  $\gamma$ . Par contre, ce domaine pourrait interagir avec l'ARNm ou un ARN ribosomal.

Dans ce contexte, il est important de noter que plusieurs mutants de la sous-unité  $\alpha$  et de la sous-unité  $\beta$  ont été sélectionnés, chez la levure, pour leur capacité à initier la traduction

du gène *HIS4* sur un codon de démarrage UUG (Donahue *et al.*, 1988 ; Cigan *et al.*, 1989). Les mutations isolées au sein d' $eIF2\alpha$  sont concentrées dans le domaine 1 de la protéine (mutants P13S et V19F d'après la numérotation de la protéine de levure) et soulignent l'importance du tonneau  $\beta$  vis-à-vis de la sélection du bon codon de démarrage (Cigan *et al.*, 1989). Le résidu proline est strictement conservé dans la famille  $e/aIF2\alpha$  et le résidu valine est présent dans 57 séquences sur les 66 séquences connues à ce jour. Pour les 9 autres séquences, la valine est remplacée par un résidu isoleucine ou leucine (8 séquences) ou par une cystéine (1 séquence). Dans la structure d' $eIF2\alpha$  de levure, les deux résidus P13 et V19 se font face, si bien que la mutation de la valine V19 en une phénylalanine pourrait déplacer le résidu proline. Ainsi, même si ces résidus ne font pas directement partie du motif de liaison des ARNs, leur mutation va très probablement affecter le repliement du tonneau  $\beta$  et par conséquent sa conformation et/ou sa stabilité (voir **Annexe 3**).

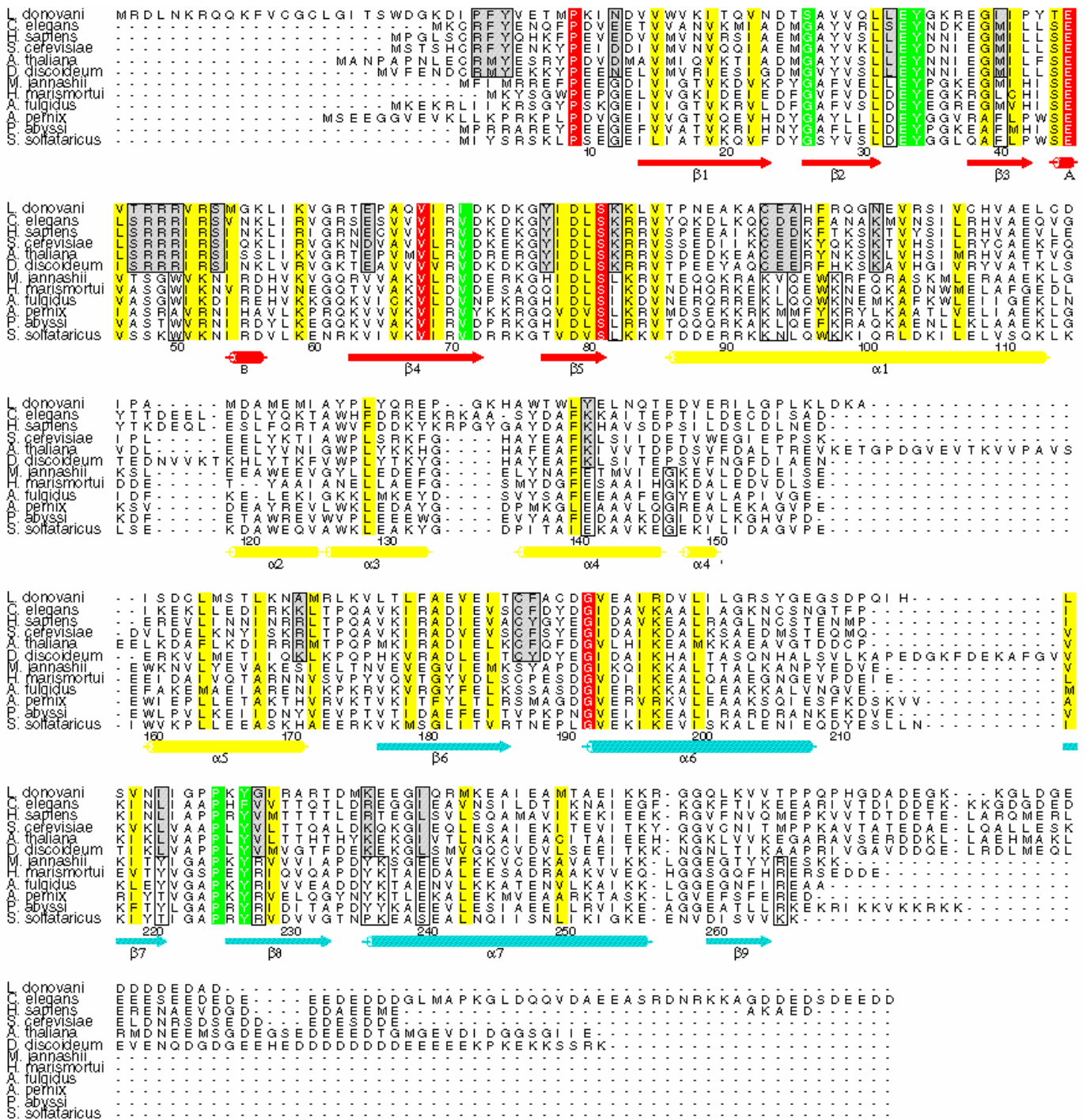
Pendant les dernières étapes de la détermination de ces structures, une équipe américaine a décrit la structure RMN de la sous-unité entière  $eIF2\alpha$  humaine (Ito *et al.*, 2004). La comparaison de la structure de  $Pa-aIF2\alpha_{D2-3}$  et de  $Pa-aIF2\alpha$  avec cette structure RMN montre que les 3 domaines d' $aIF2\alpha$  présentent une organisation et une architecture structurale similaires à celles observées pour ses homologues eucaryotiques. Le domaine 3 de notre structure peut en effet être superposé au domaine 3 de la protéine humaine avec un rmsd de 1.5 Å pour 83 C $\alpha$  comparés. La structure de la protéine humaine présente en plus une extension C-terminale riche en résidus acides. Cette extension est rencontrée chez tous les  $eIF2\alpha$  eucaryotiques (40 à 70 résidus selon les espèces), mais elle est absente chez les protéines d'Archées.

D'autre part, la mobilité que nous observons pour le domaine 3 vis-à-vis des domaines 1 et 2 est également en accord avec les enregistrements RMN obtenus pour la protéine  $eIF2\alpha$  humaine (Ito *et al.*, 2004). En effet, aucun effet NOE longue distance n'a été observé entre le domaine 3 et les domaines 1 et 2 d' $eIF2\alpha$ , ce qui est cohérent avec une absence d'interaction entre ces deux modules de la protéine. Cette mobilité pourrait permettre d'informer la sous-unité  $\gamma$  des contacts réalisés entre  $e/aIF2\alpha$  et un ARNr ou l'ARNm. Ainsi, une information qui serait transmise par ce biais pourrait être la formation de l'appariement correct entre le codon AUG de l'ARNm et l'anticodon complémentaire du Met-ARNt<sub>i</sub><sup>Met</sup>. La mobilité au sein d' $\alpha$  permettrait alors de transmettre cette information à  $e/aIF2\gamma$  et ainsi déclencher l'étape suivante de l'initiation, à savoir l'hydrolyse de la molécule de GTP portée par  $e/aIF2$ .

Finalement, compte tenu de l'homologie d' $\alpha_{D3}$  avec le domaine C-terminal de la sous-unité  $\alpha$  d' $eEF1B$  (qui est le facteur d'échange pour le facteur d'élongation  $eEF1A$ ), les auteurs proposent que le positionnement d' $\alpha_{D3}$  sur la sous-unité  $\gamma$  se fasse d'une manière identique à

celui d'eEF1B $\alpha$  sur eEF1A (Andersen *et al.*, 2001 ; Ito *et al.*, 2004). La détermination de la structure d'un complexe entre les deux sous-unités  $\alpha$  et  $\gamma$  devrait permettre de nous apporter de nouvelles informations.

L'ensemble des résultats décrits dans ce **Chapitre IV** a fait l'objet d'une publication dans *Biochemistry*.



**Figure 4.19 :** Alignement de séquences des e/aIF2 $\alpha$ . Figure dessinée avec *Alscript* (Barton, 1993). 66 séquences d'e/aIF2 $\alpha$  (43 séquences de protéines eucaryotiques et 23 séquences de protéines archéennes) ont été alignées avec *Clustal X* (Thompson *et al.*, 1997). Par souci de clarté, nous avons choisi de ne représenter que 12 séquences sur cet alignement. Cependant, la conservation des différents résidus effectuée à l'aide du programme *Plotsimilarity* a été étudiée sur l'ensemble des 66 séquences. Les résidus strictement conservés sont indiqués en blanc sur fond rouge. Les résidus très conservés (4 exceptions au maximum parmi les 66 séquences) sont indiqués en blanc sur fond vert. Les positions pour lesquelles des substitutions conservatives sont observées (2 exceptions au maximum sur l'ensemble des séquences parmi les résidus d'une même famille) sont indiquées en noir sur fond jaune. Les positions caractéristiques des séquences eucaryotiques sont encadrées dans un rectangle à fond gris et les positions caractéristiques des séquences archéennes sont encadrées dans un rectangle à fond blanc. Enfin, la position de la sérine qui est phosphorylée chez les Eucaryotes (S51 chez l'homme) est indiquée par un signe + sous les séquences. La numérotation indiquée sous les séquences correspond à celle de *P. abyssi*.

*Chapitre V : Etude  
Fonctionnelle et  
Structurale de  
l'Hétérodimère  $\alpha 1F2\alpha\gamma$*



## **Partie A : L'hétérodimère aIF2 $\alpha_{D3}\gamma$ de *P. abyssi*.**

L'étape suivante dans la compréhension du mode de reconnaissance de l'ARNt initiateur par e/aIF2 est tout naturellement l'étude de l'hétérodimère aIF2 $\alpha\gamma$  puisque cette entité est suffisante pour lier spécifiquement le Met-ARNt<sub>i</sub><sup>Met</sup>. De nombreux essais de cristallisation ont été entrepris avec diverses formes hétérodimériques du facteur archéen. C'est finalement la forme aIF2 $\alpha_{D3}\gamma$  de *P. abyssi* qui a permis d'obtenir les premiers cristaux.

### **1) Cristallisation d'aIF2 $\alpha_{D3}\gamma$ de *P. abyssi*.**

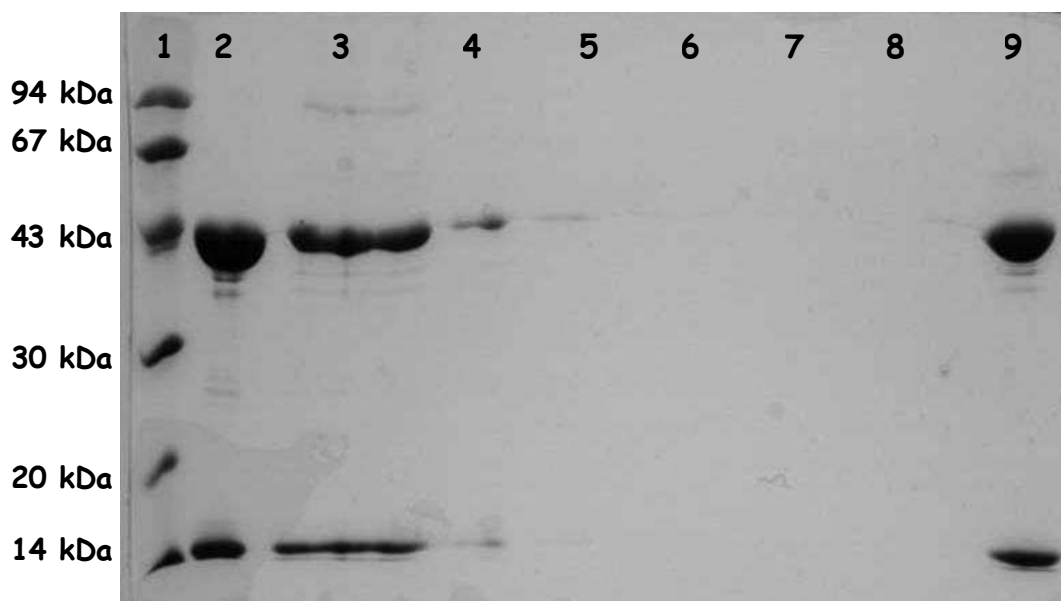
La cristallisation de l'hétérodimère aIF2 $\alpha_{D3}\gamma$  de *P. abyssi* a été réalisée manuellement à partir de la technique de la diffusion de vapeur en goutte suspendue. Les matrices d'échantillonnage commerciales fournies par Hampton Research ont été utilisées pour la recherche des conditions de cristallisation de la protéine. Les premiers essais de cristallisation ont été effectués à 24°C à partir d'une solution de protéine  $\alpha_{D3}\gamma$  fraîchement purifiée suivant le protocole décrit et concentrée à 6.6 mg.mL<sup>-1</sup>, en présence de GDPNP-Mg<sup>2+</sup> (1 mM). L'observation des gouttes a révélé la présence, au bout d'une journée, de cristaux volumineux de forme hexagonale. Ces cristaux ont été obtenus dans la condition de cristallisation n°36 de la matrice d'échantillonnage « Crystal Screen II » de Hampton (**Tableau 5.1**).

<b>Condition</b>	<b>Agent Précipitant</b>	<b>Tampon</b>
CSII - 36	4.3 M Chlorure de Sodium	0.1 M HEPES pH 7.5

**Tableau 5.1 :** Condition de cristallisation de l'hétérodimère aIF2 $\alpha_{D3}\gamma$  de *P. abyssi*.

Les cristaux ainsi obtenus sont parfaitement reproductibles dans les mêmes conditions de cristallisation. Des essais de cristallisation de l'hétérodimère ont également été effectués avec le mutant  $\gamma$ G235D. Des cristaux de taille voisine ont été obtenus dans la même condition mais ils présentent une forme singulière par rapport aux cristaux natifs (**Figure 5.3**).

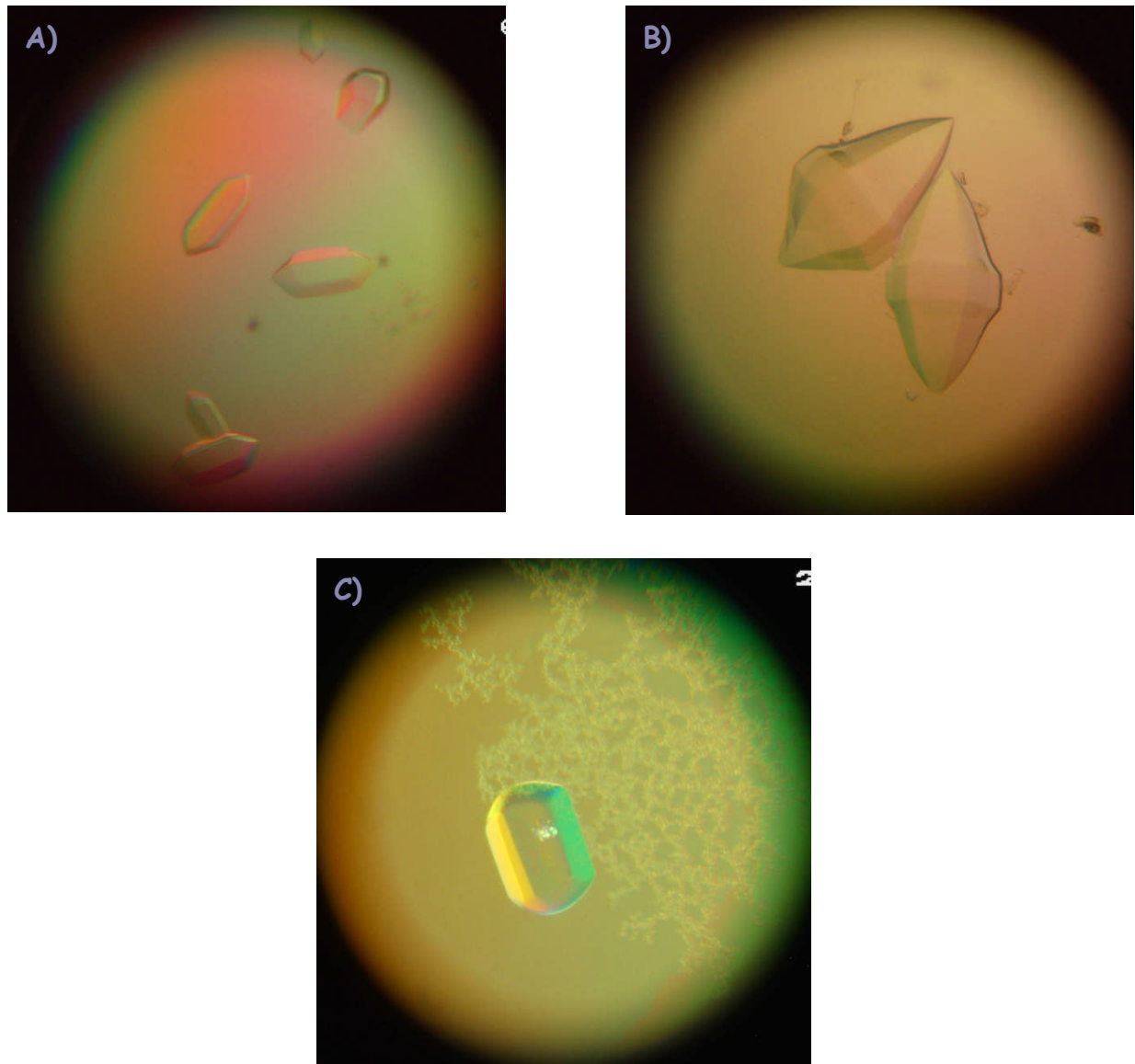
La présence dans les cristaux des deux sous unités Pa-aIF2 $\gamma$  et Pa-aIF2 $\alpha_{D3}$  a été confirmée en analysant la composition de ces cristaux par électrophorèse sur gel de polyacrylamide. Pour cela, les cristaux ont subi six lavages successifs dans des gouttes de 10  $\mu$ L de la solution réservoir. Ces étapes de lavage ont pour but d'éliminer toute trace de protéine non cristallisée qui pourrait être encore présente dans la goutte de cristallisation. Les cristaux ont finalement été dissous dans de l'eau et un aliquot de l'échantillon a été analysé par SDS-PAGE (**Figure 5.2**).



**Figure 5.2 :** Electrophorèse sur gel de polyacrylamide 12% pour analyser le contenu protéique des cristaux obtenus lors des essais de cristallisation de l'hétérodimère Pa-aIF2 $\alpha_{D3}\gamma$ . Piste 1 : marqueur de poids moléculaire ; Piste 2 : témoin  $\alpha_{D3}\gamma$  ; Piste 3 : témoin  $\alpha_{D3}\gamma$  auquel ont été ajoutés 8  $\mu$ L de la solution de cristallisation ; Piste 4 : solution restant dans la goutte de cristallisation après prélèvement des cristaux ; Piste 5 : 1<sup>ère</sup> solution de lavage, après transfert des cristaux ; Piste 6 : 2<sup>nd</sup>e solution de lavage ; Piste 7 : 5<sup>ème</sup> solution de lavage ; Piste 8 : 6<sup>ème</sup> solution de lavage ; Piste 9 : cristaux lavés et dissous dans de l'eau. Les deux protéines  $\alpha_{D3}$  et  $\gamma$  sont bien présentes dans les cristaux.

Pour l'acquisition des données, les cristaux d'aIF2 $\alpha_{D3}\gamma$  ont été trempés dans une solution contenant 4.22 M NaCl, 0.1 M HEPES pH 7.5 et 10% éthylène glycol (cryoprotectant) et immédiatement refroidis dans l'éthane liquide.

Les premiers jeux de données enregistrés pour les cristaux natifs ou mutés ont révélé un faible pouvoir diffractant : seul un jeu à 5.6 Å a pu être collecté. Au vu de cette faible résolution, une optimisation des conditions de cristallisation a été tentée à l'aide des matrices d'échantillonnages *Additives* et *Detergents Screens* d'Hampton Research. Les cristaux précédemment obtenus ont ainsi été reproduits dans la condition 4.3 M NaCl, 0.1 M HEPES pH 7.5 en ajoutant dans chaque goutte de cristallisation l'un ou l'autre des additifs/détergents de ces solutions commerciales à une dilution de 1/10<sup>ème</sup>. Les différents cristaux obtenus pour chaque additif ont subi le même traitement de cryoprotection que les cristaux initiaux. Leur pouvoir diffractant a ensuite été testé à l'ESRF. Le meilleur jeu de données a été obtenu avec l'additif HEGA-9 présent à la concentration de 39 mM dans la goutte de cristallisation (4.7 Å de résolution).



**Figure 5.3 :** Cristaux de l'hétérodimère  $\alpha\text{IF}2\alpha_{\text{D}3\gamma}$  de *P. abyssi*. **A)** Cristaux de l'hétérodimère natif obtenus en présence de 4.3 M NaCl et de 0.1 M HEPES pH 7.5. **B)** Cristaux obtenus avec l'assemblage mutant  $\alpha_{\text{D}3\gamma_{\text{G}235\text{D}}}$ . **C)** Cristaux natifs obtenus en ajoutant l'additif HEGA-9 (39 mM) dans la goutte suspendue.

## **2) Enregistrement et traitement des données.**

La collecte des jeux de données relatifs à la diffraction des rayons X par ces cristaux a été réalisée à partir des sources de rayonnement synchrotron de l'ESRF à Grenoble (ligne ID14-eh3). Un jeu de données à 4.7 Å a ainsi été collecté par la méthode de rotation du cristal. Les données de diffraction ont ensuite été traitées à l'aide des programmes *MOSFLM* (Leslie A.G.W., MRC Laboratory of Molecular Biology, Cambridge, Royaume-Uni) et *SCALA* (Collaborative Computational Project n°4, 1994). Les conditions d'acquisition pour

les cristaux natifs sont reprises dans le **Tableau 5.4** et les statistiques du jeu de données dans le **Tableau 5.5**.

	$\alpha_{D3}\gamma$ natif
Source Synchrotron	ESRF
Ligne	ID14-eh3
$\lambda$ (Å)	0.9340
Détecteur	Quantum 4R CCD
Angle d'oscillation	1°
Distance au détecteur (mm)	300
Temps d'exposition (s)	60 (3 passages)
Nombre de clichés	30

**Tableau 5.4 :** Conditions d'acquisition des données de diffraction des cristaux natifs d'aIF2 $\alpha_{D3}\gamma$  de *P. abyssi*.

Cristal	$\alpha_{D2-3}$ natif
Groupe d'espace	P3 <sub>2</sub> 21
Paramètres de maille	a = b = 124.76 Å c = 139.06 Å
Résolution (Å)	4.7
Réflexions uniques	5891
Complétude (%)	88.2
Redondance	2.0
R <sub>sym</sub> (I) (%) <sup>1</sup>	5.1 (32.3)

**Tableau 5.5 :** Statistiques du traitement des données. <sup>1</sup> Pour le R<sub>sym</sub>, les chiffres entre parenthèses correspondent aux statistiques dans la plus haute tranche de résolution.

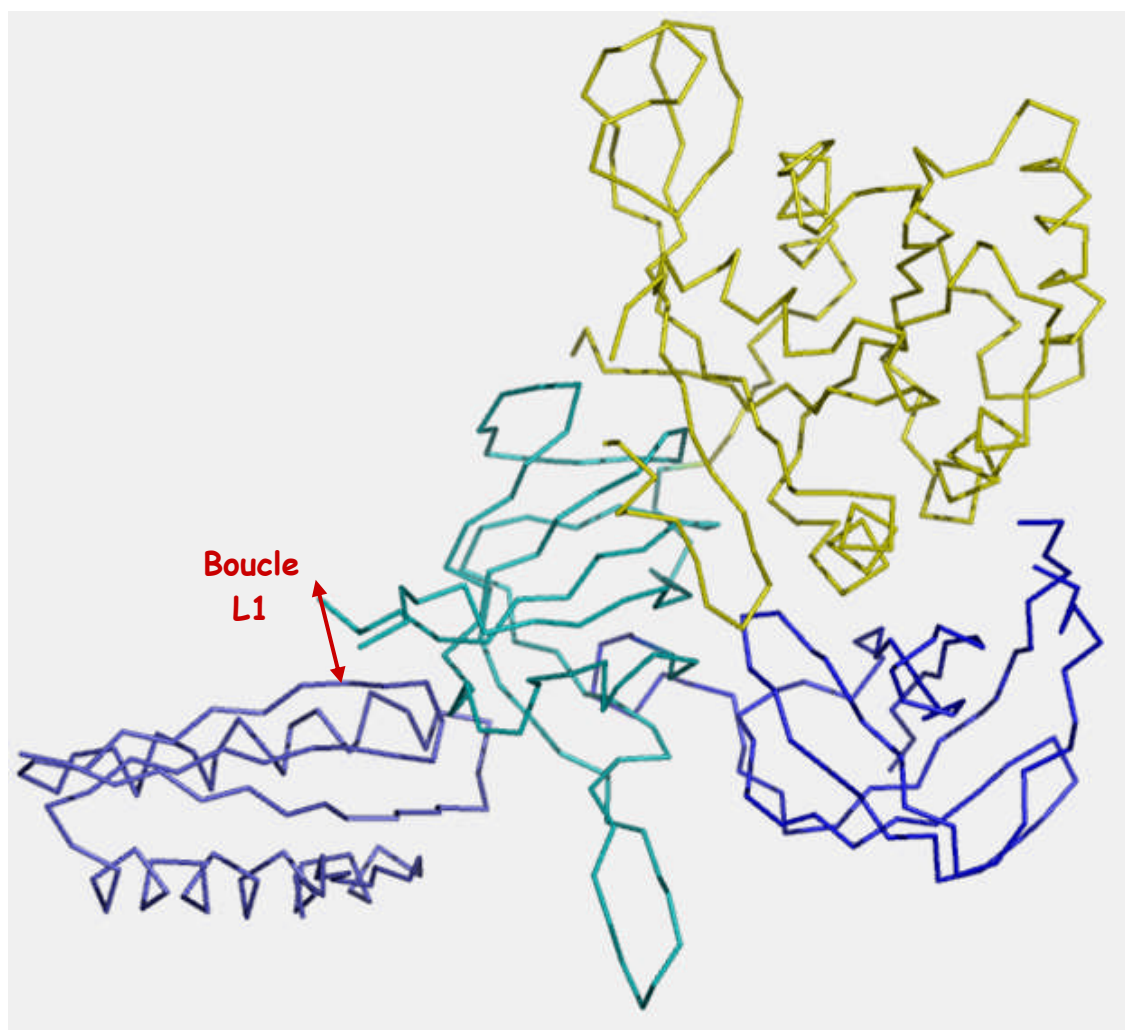
### **3) Tentatives pour résoudre la structure.**

Malgré la faible résolution de ce jeu de données, une tentative de phasage a été entreprise par remplacement moléculaire à l'aide du programme *PHASER* (Storoni *et al.*, 2004). Les structures des protéines aIF2 $\gamma$  et aIF2 $\alpha_{D2-3}$  de *P. abyssi* ont servi de modèles pour résoudre cette structure.

Dans un premier, les domaines 1, 2 puis 3 d'aIF2 $\gamma$  de *P. abyssi* ont été recherchés séparément. Une solution unique a été obtenue dans chaque cas (Zscores respectifs de 9.9, 14.5 et 19.5). Le domaine 3 d'aIF2 $\alpha$ , basé sur la structure d'aIF2 $\alpha_{D2-3}$  de *P. abyssi*, a ensuite été identifié sous forme d'une solution également unique avec un Zscore de 6.3. Les solutions uniques relatives aux trois domaines de  $\gamma$  et au domaine C-terminal d' $\alpha$  fournies par le programme *PHASER* ont alors permis de construire une première carte de densité électronique. Celle-ci a été améliorée par lissage de solvant à l'aide du programme *DM* (Cowtan and Main, 1993) du groupe de programmes CCP4 (Collaborative Computational Project N°4, 1994). Le modèle a alors été affiné par une minimisation d'énergie en considérant les trois domaines de  $\gamma$  et le domaine 3 d' $\alpha$  comme des ensembles rigides à l'aide du programme *CNS* (Brunger *et al.*, 1998). Un facteur R de 40 % et un facteur R<sub>libre</sub> de 42% ont ainsi été obtenus. Les tentatives de remplacement manuel du modèle dans la densité électronique et les différents tours d'affinement réalisés n'ont pas permis d'améliorer ces paramètres du fait de la basse résolution.

Le modèle non affiné obtenu par remplacement moléculaire est représenté sur la **Figure 5.6**. Bien qu'aucun affinement n'ait pu être réalisé sur ce modèle initial, celui-ci est en accord avec les données biochimiques décrites dans le **Chapitre III** de ce manuscrit. En effet, le domaine 3 d'aIF2 $\alpha$  se retrouve bien à proximité de la sous-unité  $\gamma$ , et notamment, il est au contact de la boucle L1 (désordonnée dans la structure modèle d'aIF2 $\gamma$  de *P. abyssi* et donc non visualisable) par laquelle aIF2 $\gamma$  est censé interagir avec aIF2 $\alpha$ .

Ce modèle propose donc une piste particulièrement intéressante et relativement fiable - car corroborée par les données expérimentales - du mode d'interaction des deux sous-unités d'aIF2 impliquées dans la reconnaissance du substrat ARNt. Cependant, la résolution trop basse de ce jeu de données ne permet pas d'obtenir une structure définitive du complexe. Malgré divers essais, aucune amélioration du pouvoir diffractant de ces cristaux n'a pu être obtenue et aucun autre assemblage hétérodimérique n'a donné lieu à une cristallisation conjointe des deux sous-unités. Nous avons donc choisi de purifier les sous-unités d'aIF2 en utilisant les gènes d'une autre archée.



**Figure 5.6 :** Modèle rigide de l'hétérodimère  $\alpha\text{IF}2\alpha_{\Delta 3}\gamma$  de *P. abyssi* obtenu par remplacement moléculaire à 4.7 Å de résolution.

## **Partie B : Caractérisation biochimique du facteur aIF2 de *Sulfolobus solfataricus*.**

### **1) Introduction.**

Nous avons choisi comme nouvelle cible le facteur aIF2 de *S. solfataricus*. Le séquençage du génome complet de *S. solfataricus* (She *et al.*, 2001) a permis d'identifier les gènes codant pour les trois sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$  du facteur de démarrage : SSO1050, SSO2381 et SSO0412.

Les séquences des protéines aIF2 issues de *P. abyssi* et *S. solfataricus* sont bien conservées. Ainsi, les sous-unités aIF2 $\gamma$  sont homologues à 56%. Les taux sont un peu plus faibles pour  $\alpha$  et  $\beta$  (respectivement 35 et 30% d'identité de séquence). Le **Tableau 5.7** reprend les caractéristiques des différentes protéines de l'hétérotrimère aIF2 de *Sulfolobus solfataricus*.

<b>Protéine</b>	<b>Poids moléculaire</b>	<b>Point isoélectrique</b>	<b>Facteur d'absorption molaire en L.cm<sup>-1</sup>.mol<sup>-1</sup></b>
aIF2 $\alpha$	29.7 kDa	8.93	42.8 $\times 10^3$
aIF2 $\beta$	15.9 kDa	8.80	12.7 $\times 10^3$
aIF2 $\gamma$	45.8 kDa	8.52	48.4 $\times 10^3$

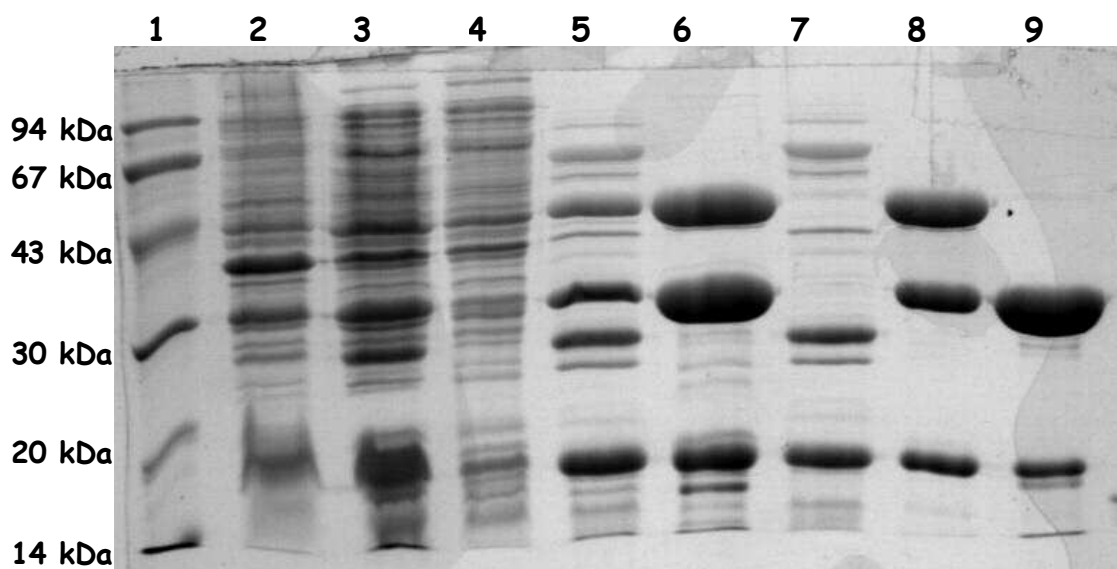
**Tableau 5.7:** Caractéristiques générales des différentes sous-unités du facteur aIF2 de *S. solfataricus* déterminées à l'aide du programme *DNAid* (Dardel et Bensoussan, 1988).

### **2) Clonage des gènes codant pour les 3 sous-unités d'aIF2 de *S. solfataricus* et purification des protéines.**

Le gène codant pour aIF2 $\alpha$  de *S. solfataricus* a été cloné au sein du vecteur pET3alpa, déjà décrit dans le **Chapitre III**, entre les sites de restriction *NdeI* et *XhoI* du fragment additionnel. Les gènes codant pour aIF2 $\gamma$  et aIF2 $\beta$  *S. solfataricus* ont été clonés dans le vecteur pET15blpa, entre les sites de restriction *NcoI* et *SalI* pour  $\gamma$  (ce clonage enlève le tag His<sub>6</sub> N-terminal) et entre les sites *NcoI* et *XhoI* pour  $\beta$ . Les deux formes tronquées d'aIF2 $\alpha$ , correspondant respectivement au domaine 3 ( $\alpha_{D3}$ ) et aux domaines 2 et 3 ( $\alpha_{D2-3}$ ), ont été obtenues par mutagénèse dirigée par PCR en introduisant un codon de démarrage au niveau des résidus Glu174 et Asp88 respectivement. Les protéines codées par ces gènes ont été produites dans la bactérie *Escherichia coli* Rosetta (DE3) pLacIRARE, déjà décrite au

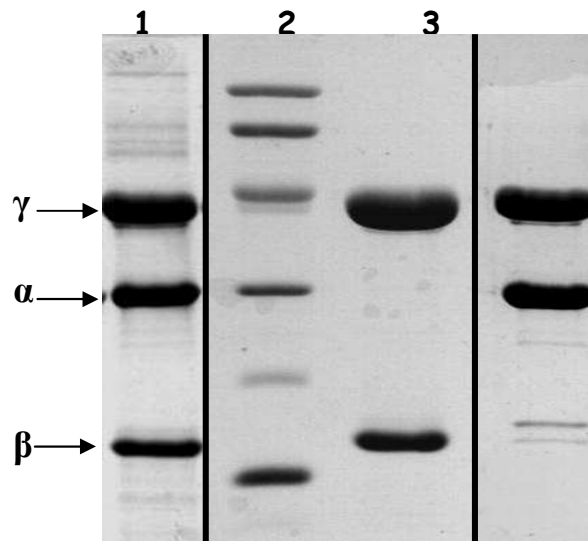
**Chapitre III.** La surproduction des protéines d'intérêt est déclenchée en ajoutant de l'IPTG (1 mM) dans le milieu de culture lorsque la densité optique à 650 nm atteint 1,6. Ces bactéries se sont révélées particulièrement efficaces pour la surproduction des protéines de Ss-aIF2.

Le mode de purification des monomères et des différents assemblages de Ss-aIF2 est décrit dans la partie **Matériel et Méthodes**. Brièvement, la provenance de ces protéines a une nouvelle fois été utilisée : un chauffage limité à 65°C de l'extrait protéique brut obtenu après lyse cellulaire a permis la dénaturation de la majeure partie des protéines bactériennes tandis que les protéines de l'hyperthermophile *S. solfataricus* restaient stables (un chauffage prolongé à 80°C dénature les protéines de *S. solfataricus*). D'autre part, dans le cas de *S. solfataricus*, tous les monomères (à l'exception de Ss-aIF2 $\gamma$ ) et assemblages du facteur ont pu être sélectivement accrochés sur une colonne échangeuse de cations à basse force ionique. Cette propriété a permis de raccourcir considérablement le processus de purification. Enfin, lors de la purification de préparations contenant la sous-unité  $\gamma$ , du GTP-Mg<sup>2+</sup> a systématiquement été ajouté à chaque étape de purification (lors de la lyse cellulaire par sonication puis après chaque colonne chromatographique) dans l'échantillon protéique pour éviter la protéolyse de  $\gamma$  en particulier au niveau de la boucle GKT impliquée dans la liaison des nucléotides guanines. La **Figure 5.8** retrace les étapes de purification de l'hétérotrimère natif Ss-aIF2.



**Figure 5.8 :** Purification de l'hétérotrimère aIF2 de *Sulfolobus solfataricus*. Des échantillons de 20  $\mu$ L correspondant aux différentes étapes de purification ont été déposés sur gel dénaturant de polyacrylamide 12%. Piste 1 : marqueur de poids moléculaire ; Piste 2 : culot obtenu après lyse des cellules d'*E. coli* par sonication ; Piste 3 : surnageant obtenu après lyse des cellules bactériennes par sonication ; Piste 4 : culot obtenu après chauffage de l'extrait protéique à 65°C ; Piste 5 : surnageant obtenu après chauffage de l'extrait protéique à 65°C ; Piste 6 : élution de la S-Hiload ; Piste 7 : fractions de tête de la S-Hiload ; Piste 8 : pic de la Superdex 200 correspondant à l'hétérotrimère Ss-aIF2 ; Piste 9 : pic de la Superdex 200 correspondant à l'excès d' $\alpha$  et de  $\beta$  (sur ce tamis moléculaire, les deux sous-unités sont éluées avec des temps de rétention voisins).

Les protéines d'aIF2 de *S. solfataricus* présentent une homogénéité supérieure à celle des protéines de *P. abyssi*, comme en témoigne la **Figure 5.9**.



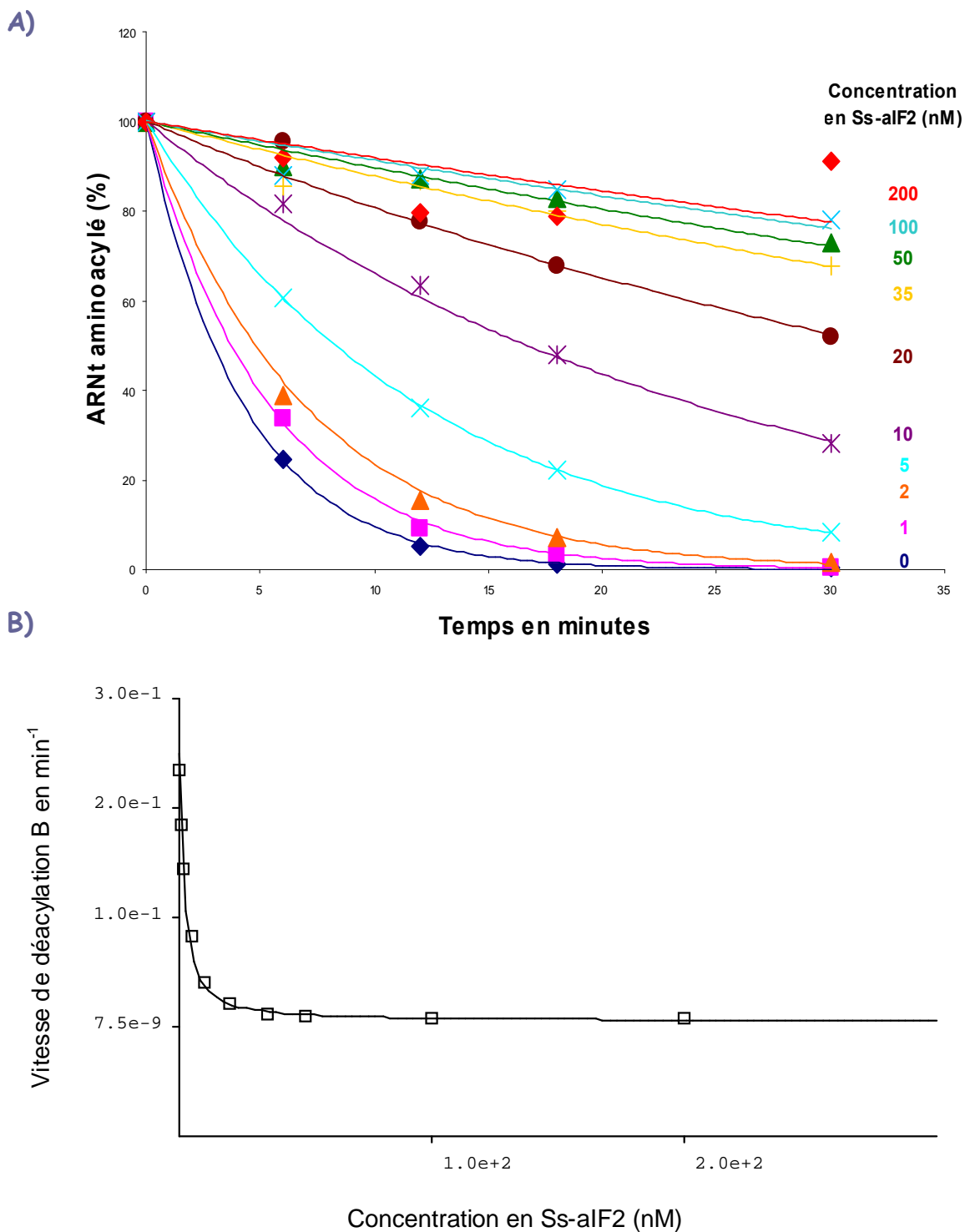
**Figure 5.9 :** migration électrophorétique de différents assemblages des 3 sous-unités purifiées de Ss-aIF2. Piste 1 : hétérotrimère Ss-aIF2 $\alpha\beta\gamma$  (46, 30 et 16 kDa) ; Piste 2 : marqueur de poids moléculaire ; Piste 3 : hétérodimère Ss-aIF2 $\beta\gamma$ ; Piste 4 : hétérodimère Ss-aIF2 $\alpha\gamma$ .

### 3) Affinité vis-à-vis du Met-ARNt<sub>i</sub><sup>Met</sup> et rôle des différentes sous-unités.

Avant de passer à l'étude structurale du complexe, nous avons souhaité vérifier les résultats obtenus avec l'hétérotrimère aIF2 de *P. abyssi* concernant le rôle joué par chaque sous-unité dans la liaison à l'ARNt initiateur. Pour cela, le test de mesure de l'affinité des différents assemblages d'aIF2 pour l'ARNt a été reproduit avec les protéines de *S. solfataricus*, dans des conditions identiques à celles décrites au **Chapitre III**. Les premiers résultats obtenus sur l'hétérotrimère suggéraient une meilleure affinité pour le Met-ARNt<sub>i</sub><sup>Met</sup> que dans le cas de Pa-aIF2 (d'au moins un facteur 10). Nous avons donc reproduit les mesures en nous plaçant cette fois-ci à 2 nM en Met-ARNt dans le test, pour permettre une mesure fiable de ces constantes de dissociation plus basses (voir **Chapitre III** pour les conditions du test).

Une fois encore, le Met-ARNt<sub>i</sub><sup>Met</sup> d'*E. coli* a été utilisé dans le test. En effet, l'ARNt initiateur de *S. solfataricus* est très proche de celui de *P. abyssi*, du point de vue de sa séquence nucléotidique, et par conséquent de l'ARNt initiateur bactérien. Les constantes de dissociation des différents complexes Ss-aIF2:GDPNP:ARNt<sub>i</sub> ont été mesurées à 51°C en présence du Met-ARNt<sub>i</sub><sup>Met</sup> d'*E. coli* aminoacylé avec de la <sup>35</sup>S-Met. Dans le cas de *S. solfataricus*, les assemblages hétérodimériques et hétérotrimériques ont été directement purifiés avant utilisation dans le test de protection (contrairement au cas de *P. abyssi* où les

assemblages avaient été réalisés en mélangeant les différentes sous-unités dans un rapport stoechiométrique).



**Figure 5.10 :** Mesure de la constante de dissociation du Met-ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* pour l'hétérotrimère Ss-aIF2. A) Cinétique de déacylation du Met-ARNt<sub>f</sub><sup>Met</sup> en présence de concentrations croissantes du facteur Ss-aIF2. Chaque courbe de déacylation est modélisée pour une exponentielle simple. B) Tracé des vitesses de déacylation en fonction de la concentration initiale en aIF2. Le programme MC-Fit permet de déduire de cette courbe la valeur du Kd pour le complexe considéré.

Le **Tableau 5.11** reprend les valeurs de Kd obtenues pour les complexes aIF2 de *S. solfataricus*.

	$\alpha$	$\beta$	$\gamma$	Kd (nM)
1	×	×	×	1.5 $\pm$ 0.5
2			×	> 100 000
3	×		×	7.5 $\pm$ 1.0
4	$\alpha_{D3}$		×	7.3 $\pm$ 1.0

**Tableau 5.11 :** Constantes de dissociation des assemblages d'aIF2 de *S. solfataricus* mentionnés en présence du Met-ARNt<sup>Met</sup> d'*E. coli* 2 nM et de 1 mM GDPNP-Mg<sup>2+</sup>. Les sous-unités présentes dans l'assemblage testé sont indiquées par une croix.

En présence de l'hétérotrimère natif, une constante de dissociation de 1.5 nM est obtenue à 51°C (soit un facteur 60 d'augmentation de l'affinité comparativement à l'hétérotrimère de *P. abyssi*). Globalement, tous les Kds mesurés avec les protéines de *S. solfataricus* sont améliorés par rapport à ceux obtenus pour *P. abyssi*. Ceci pourrait être dû aux différences de conditions de croissance des deux organismes. En effet, la température optimale de croissance de *P. abyssi* est de 80°C tandis que celle de *S. solfataricus* n'est que de 65°C. La température utilisée dans ce test, 51°C, est donc plus proche de l'optimum de *S. solfataricus*.

En présence de l'hétérodimère  $\alpha\gamma$ , un Kd de 7.5 nM est obtenu. L'affinité est réduite d'un facteur 5 par rapport à celle mesurée pour l'hétérotrimère natif. Dans le cas de Ss-aIF2, la sous-unité  $\beta$  pourrait contribuer à la liaison de l'ARNt initiateur. Cette contribution reste toutefois très faible en comparaison avec celle de  $\alpha$  (gain supérieur à 10<sup>4</sup>). De plus, en présence du domaine C-terminal d' $\alpha$  et de la sous-unité  $\gamma$ , une affinité similaire à celle de l'hétérodimère natif  $\alpha\gamma$  est retrouvée. Ainsi, comme chez *P. abyssi*,  $\alpha$  est essentiel à la liaison optimale de l'ARNt initiateur méthionylé et au sein de Ss-aIF2 $\alpha$ , le domaine 3 à lui seul peut accomplir cette fonction.

Cependant, contrairement au cas de *P. abyssi*, la sous-unité  $\gamma$  seule ne semble pas capable de lier l'ARNt ou tout au moins avec une affinité très réduite. En effet, aucune protection contre la déacylation n'est observée même en présence de 100  $\mu$ M en Ss-aIF2 $\gamma$ . Dans le cas d'aIF2 $\gamma$  de *P. abyssi*, une constante de dissociation de 5  $\mu$ M avait pu être mesurée et elle avait été retrouvée pour le dimère  $\beta\gamma$ . La différence de comportement des sous-unités  $\gamma$  de ces deux organismes est surprenante et reste pour le moment inexplicée. Elle pourrait être à nouveau due aux conditions de croissance qui diffèrent entre ces deux organismes.

## **Partie C : Résolution de la structure de l'hétérodimère aIF2 $\alpha\gamma$ de l'archée *Sulfolobus solfataricus*.**

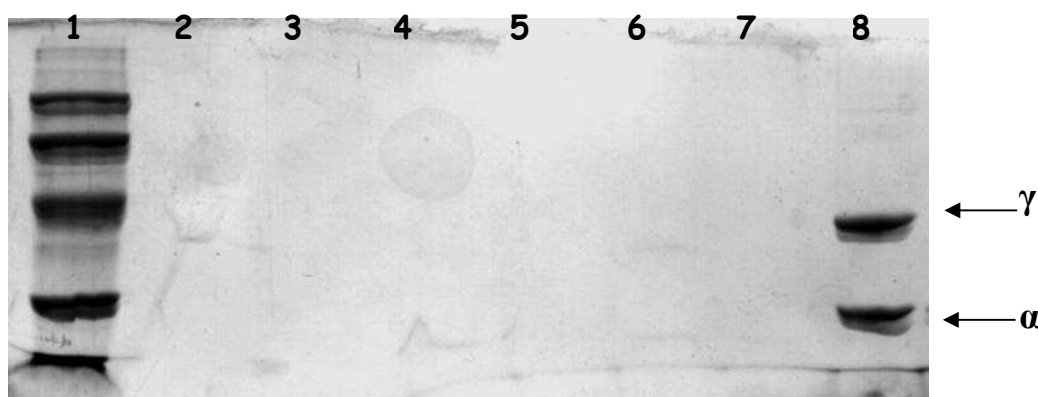
### **1) Cristallisation de Ss-aIF2 $\alpha\gamma$ et enregistrement des données.**

La cristallisation de l'hétérodimère Ss-aIF2 $\alpha\gamma$  a été réalisée à partir de la technique de la diffusion de vapeur en goutte assistée à l'aide d'un robot *Genesis RSP100* (TECAN). Les matrices d'échantillonnage commerciales fournies par Hampton Research ont été utilisées pour la recherche des conditions de cristallisation de la protéine. Les premiers essais de cristallisation ont été effectués à 24°C à partir d'une solution protéique fraîchement préparée suivant le protocole décrit. Cette solution protéique contient l'hétérodimère Ss-aIF2 $\alpha\gamma$  à 9 mg.mL<sup>-1</sup> ainsi que du GDPNP-Mg<sup>2+</sup> à 1 mM. Des gouttes de 2  $\mu$ L ont été réalisées (1  $\mu$ L de solution protéique et 1  $\mu$ L de solution contenant l'agent précipitant). Après une semaine, l'observation des gouttes a révélé la présence de cristaux en forme d'aiguilles, ces derniers croissant autour de noyaux de nucléation (**Figure 5.14**). Les cristaux d'aIF2 $\alpha\gamma$  ont été obtenus en utilisant les deux solutions précipitantes décrites dans le **Tableau 5.12**.

Condition	Agent Précipitant	Additif
<i>Grid Screen PEG 6000 - 10</i>	10% PEG 6000	0.1 M HEPES pH 7.0
<i>Grid Screen PEG 6000 - 12</i>	10% PEG 6000	0.1 M Bicine pH 9.0

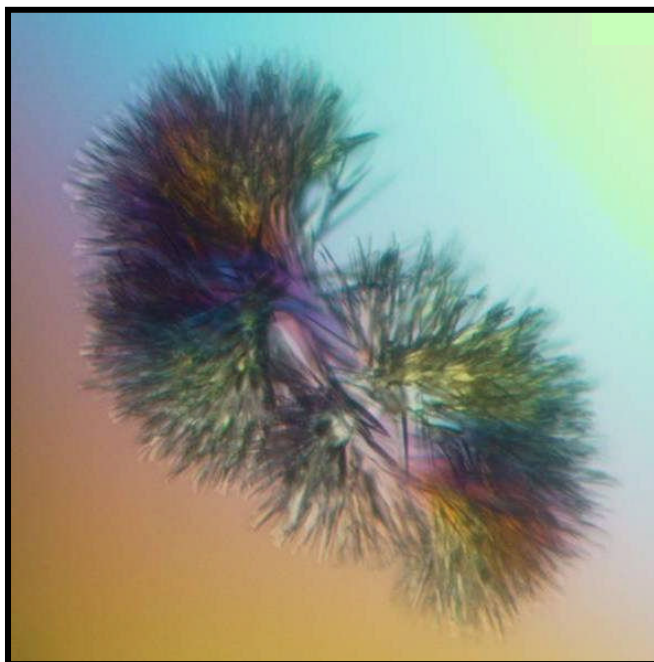
**Tableau 5.12:** Conditions de cristallisation du dimère aIF2 $\alpha\gamma$  de *S. solfataricus*.

Le contenu protéique de ces cristaux a été analysé sur gel d'électrophorèse en conditions dénaturantes. Les cristaux ont été soumis à 6 lavages successifs dans la solution réservoir puis ils ont été dissous dans de l'eau (**Figure 5.13**).



**Figure 5.13 :** Electrophorèse sur gel de polyacrylamide 12% pour analyser le contenu protéique des cristaux obtenus lors des essais de cristallisation de l'hétérodimère Ss-aIF2 $\alpha\gamma$ . Piste 1 : marqueur de poids moléculaire ; Pistes 2 à 7 : solutions de lavage 1 à 6, après transfert des cristaux ; Piste 8 : cristaux lavés et dissous dans de l'eau. Les deux protéines  $\alpha$  et  $\gamma$  sont bien présentes dans les cristaux.

Ces cristaux se sont révélés peu stables dans l'éthane liquide après cryogénéisation. Ils ont donc été directement montés dans les boucles lors de l'acquisition des données. Au préalable, ils ont été trempés dans une solution contenant la solution précipitante et 20% d'éthylène glycol comme cryoprotectant. Puis ils ont été immédiatement refroidis dans le flux d'azote du montage, avant de débiter l'enregistrement des données de diffraction.



**Figure 5.14 :** Cristaux du dimère natif aIF2 $\alpha\gamma$  de *S. solfataricus* obtenus dans la condition de cristallisation 10% PEG 6000, 0.1 M HEPES pH 7.0.

Les données de diffraction relatives aux cristaux de Ss-aIF2 $\alpha\gamma$  ont été collectées à 100K à l'ESRF (Grenoble) sur la ligne ID14-eh1. Un jeu de données à 3.0 Å de résolution a été enregistré. Les conditions d'acquisition pour les cristaux natifs sont reprises dans le **Tableau 5.15**.

	$\alpha\gamma$ natif
Source Synchrotron	ESRF
Ligne	ID14-eh1
$\lambda$ (Å)	0.9340
Détecteur	Quantum 4R CCD
Angle d'oscillation	0.5°
Distance au détecteur (mm)	250
Temps d'exposition (s)	10 (3 passages)
Nombre de clichés	180

**Tableau 5.15 :** Conditions d'acquisition des données de diffraction des cristaux natifs de Ss-aIF2 $\alpha\gamma$ .

Les cristaux d' $\alpha\gamma$  présentent une forte mosaïcité et de grands paramètres de maille. Pour pallier ce problème, un angle d'oscillation assez faible ( $0.5^\circ$ ) a été utilisé au cours de l'enregistrement. Malgré cela, l'utilisation du logiciel *MOSFLM* (Leslie A.G.W., MRC Laboratory of Molecular Biology, Cambridge, Royaume-Uni) n'a pas permis de traiter correctement les clichés de diffraction. Par conséquent, le jeu de données à  $3.0 \text{ \AA}$  relatif à  $\alpha\gamma$  a traité à l'aide du programme *XDS* (Kabsch, 1988). Ce logiciel traite les taches de diffraction en trois dimensions contrairement à *MOSFLM* qui ne les traite que cliché par cliché, donc en deux dimensions. *XDS* est donc plus approprié pour traiter des images où les taches de diffraction sont très chevauchantes, du fait de la forte mosaïcité. Les statistiques concernant le traitement des données sont reprise dans le **Tableau 5.16**.

Cristal	$\alpha$ D2-3 natif
Groupe d'espace	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Paramètres de maille	a = 39.6 Å
	b = 112.9 Å
	c = 194.1 Å
Résolution (Å)	3.0
Réflexions uniques	17816
Complétude (%)	97.5
Redondance	3.6
Rsym(I) (%) <sup>1</sup>	7.4 (30.2)

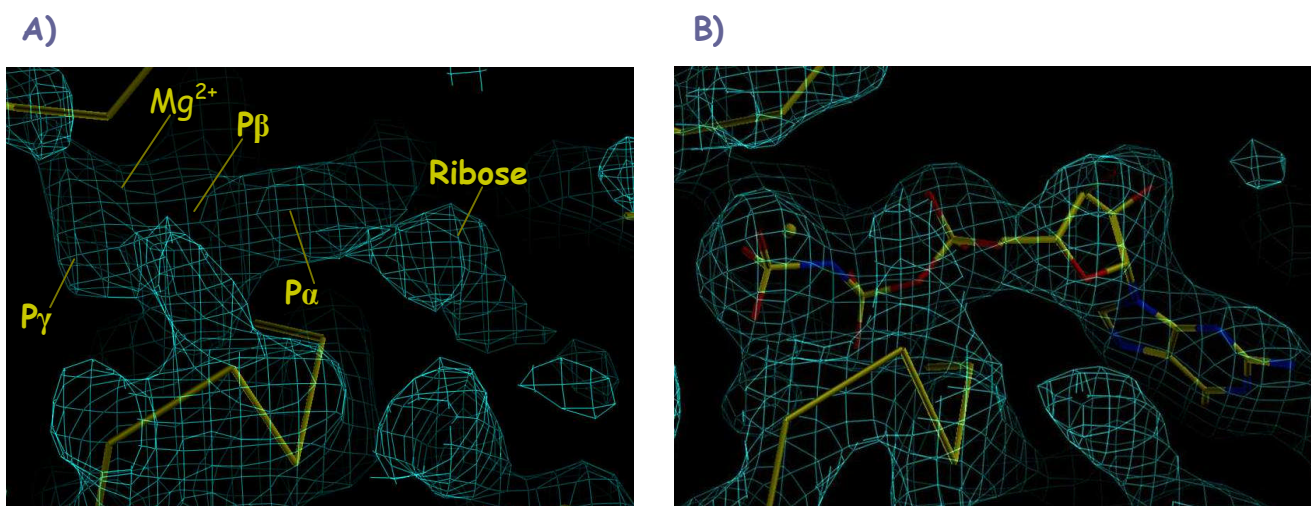
**Tableau 5.16 :** Statistiques du traitement des données. <sup>1</sup> Pour le Rsym, les chiffres entre parenthèses correspondent aux statistiques dans la plus haute tranche de résolution.

## 2) Résolution de la structure et affinement du modèle.

La recherche des phases pour le jeu de données à  $3.0 \text{ \AA}$  de Ss-aIF2 $\alpha\gamma$  a ensuite été entreprise par remplacement moléculaire à l'aide du programme *PHASER* (Storoni *et al.*, 2004). Les trois domaines d'aIF2 $\gamma$  ont d'abord été recherchés à l'aide du modèle structural fourni par la structure d'aIF2 $\gamma$  de *P. abyssi* (Schmitt *et al.*, 2002). Trois solutions uniques ont été successivement obtenues pour les domaines 1, 2 et 3 d'aIF2 $\gamma$  avec les Zscores respectifs de 7.6, 11.6 et 14.1. La seconde étape a consisté à rechercher la position des domaines d'aIF2 $\alpha$  au sein de la maille cristalline. Pour cela, différents modèles structuraux ont été utilisés : aIF2 $\alpha$ <sub>D2-3</sub> et aIF2 $\alpha$  de *P. abyssi* précédemment résolues, eIF2 $\alpha$ <sub>D1-2</sub> de levure (Dhaliwal and Hoffman, 2003), eIF2 $\alpha$ <sub>D1-2</sub> humain (Nonato *et al.*, 2002) et eIF2 $\alpha$  humain (Ito *et al.*, 2004). Dans chaque cas, les solutions ont été recherchées en utilisant soit le modèle structural entier, soit les domaines séparés de chaque protéine. Différentes combinaisons ont été utilisées pour l'ordre de recherche. Des modèles structuraux correspondant à ces protéines

virtuellement mutées en chaînes polyalanines ont également été utilisés pour faciliter la recherche de solutions. Cependant, aucune solution pour la sous-unité  $\alpha$  ou l'un de ses domaines séparés n'a pu être obtenue par remplacement moléculaire, quel que soit le modèle structural utilisé.

La structure du complexe aIF2 $\alpha\gamma$  de *S. solfataricus* a alors été résolue comme suit : les trois solutions uniques obtenues pour les trois domaines séparés d'aIF2 $\gamma$  ont permis de calculer une première carte de densité à 3.0 Å de résolution. Le modèle structural obtenu par remplacement moléculaire a alors été positionné dans cette carte et les résidus non communs entre aIF2 $\gamma$  de *S. solfataricus* et aIF2 $\gamma$  de *P. abyssi* ont été mutés en alanine (56% d'identité de séquence entre les deux protéines). La carte de densité initiale a alors été améliorée par lissage de solvant à l'aide du programme *DM* (Cowtan and Main, 1993).



**Figure 5.17 :** amélioration de la densité autour du nucléotide GDPNP et de l'ion Mg<sup>2+</sup> au cours de l'affinement du modèle pour Ss-aIF2 $\alpha\gamma$ . **A)** Carte de densité électronique après 7 tours d'affinement : seule la molécule  $\gamma$  est placée dans le modèle structural à ce stade. **B)** Carte de densité électronique finale, après reconstruction de tout le modèle  $\alpha\gamma$ .

Au sein de cette carte, de la densité était très clairement présente à l'extérieur de la molécule d'aIF2 $\gamma$  et de ses symétriques, suggérant de la place pour une autre molécule au sein de la maille cristalline (une molécule d' $\alpha\gamma$  par unité asymétrique). De plus, cette densité s'organisait notamment autour du domaine 2 d'aIF2 $\gamma$ , site de liaison présumé du domaine 3 d'aIF2 $\alpha$ . La densité additionnelle n'était cependant pas suffisamment organisée pour permettre la reconnaissance d'un quelconque motif structural appartenant à aIF2 $\alpha$ . Le modèle d'aIF2 $\gamma$  a alors été affiné avec différents cycles de construction manuelle et de minimisation d'énergie à l'aide des programmes *O* (Jones *et al.*, 1991) et *CNS* (Brunger *et al.*, 1998) en utilisant seulement les données comprises entre 12 et 3.0 Å et tronquées à 2 $\sigma$ . Ce premier affinement a notamment permis de retracer les chaînes latérales manquantes au sein d'aIF2 $\gamma$  de *S. solfataricus* et de repositionner les carbones C $\alpha$  par rapport au modèle aIF2 $\gamma$  de *P.*

*abyssi*. De plus, au sein du domaine G de Ss-aIF2 $\gamma$ , la densité du GDPNP et des ions Mg<sup>2+</sup> et Zn<sup>2+</sup> est apparue, ce qui constitue un signe de bon affinement (**Figure 5.17**).

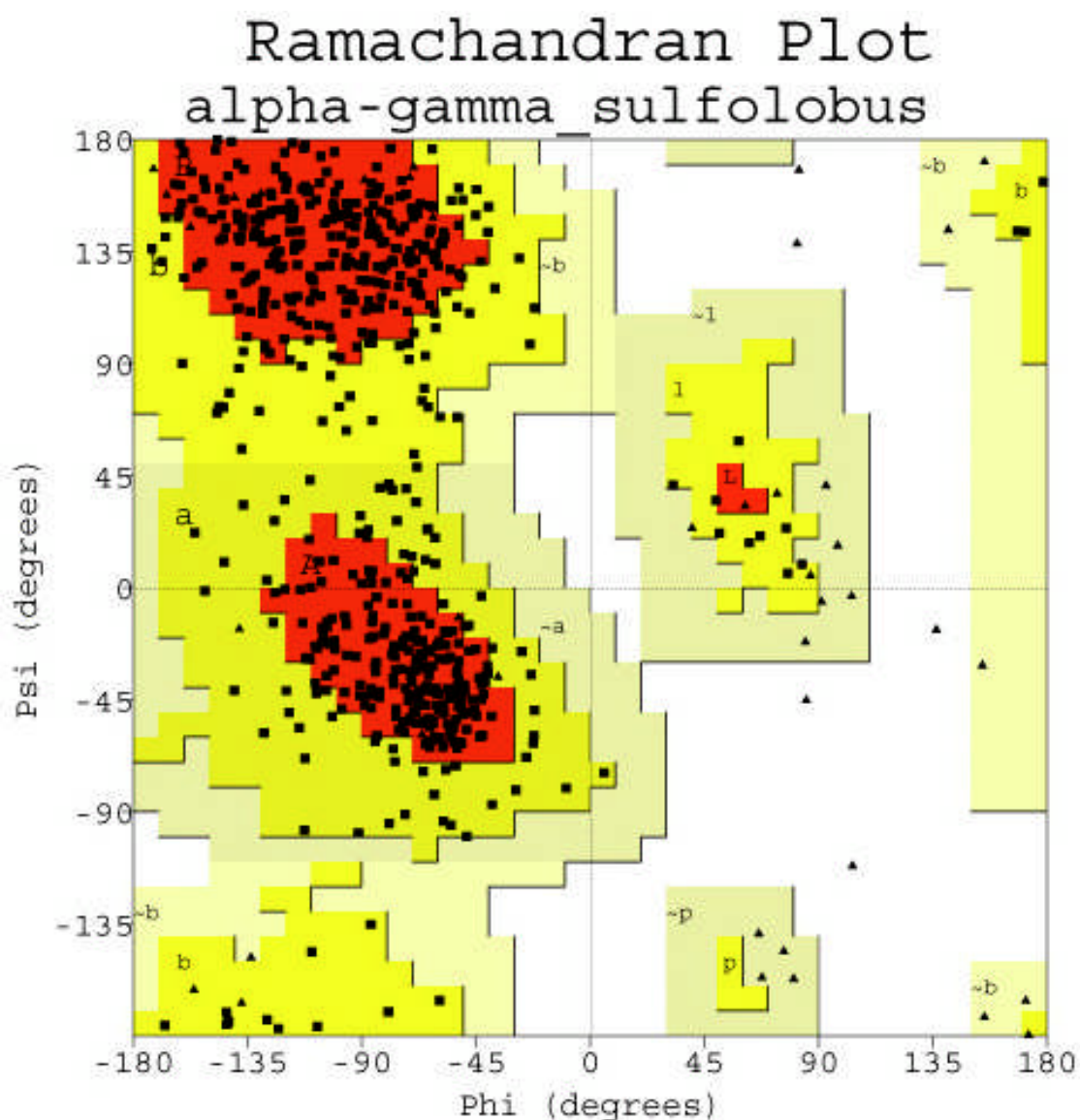
Cette première étape a permis d'obtenir une structuration de la densité électronique autour du domaine 2 de Ss-aIF2 $\gamma$ . Des éléments de structure caractéristiques du domaine 3 d'aIF2 $\alpha$  ont été identifiés et ont permis le positionnement de ce domaine dans le modèle. Là encore, les résidus divergents entre aIF2 $\alpha$  de *S. solfataricus* et aIF2 $\alpha$  de *P. abyssi* ont été changés en alanine dans un premier temps. Puis les résidus corrects ont été réinsérés au fur et à mesure des tours d'affinements. Finalement, les domaines 2 puis 1 d'aIF2 $\alpha$  de *P. abyssi* ont pu être positionnés sans ambiguïté dans la densité électronique. En fin d'affinement, le modèle a été complété avec plusieurs molécules d'eau et amélioré en affinant les valeurs du facteur B par groupes d'atomes. Le **Tableau 5.18** reprend ces différents paramètres. Le modèle obtenu pour le dimère aIF2 $\alpha\gamma$  de *S. solfataricus* présente un facteur R final de 23.2% et un facteur R<sub>libre</sub> de 29.7% à 3.0 Å de résolution.

	aIF2 $\alpha\gamma$ de <i>S.solfataricus</i>
R / R <sub>libre</sub> (calculé sur 6% des réflexions)	22.4 / 29.1
Rmsd liaisons (Å) / Rmsd angles (°)	0.0093 / 1.7989
Nombre de résidus	678
Nombre de molécules d'eau	42
<b>Facteurs B</b>	
Protéine $\alpha$ (Å <sup>2</sup> )	59.39
Protéine $\gamma$ (Å <sup>2</sup> )	40.11
Molécules d'eau (Å <sup>2</sup> )	33.19
<b>Diagramme de Ramachandran</b>	
Résidus dans la zone favorable (%)	76.6
Résidus dans les zones tolérées (%)	23.4
Résidus dans les zones généreusement tolérées (%)	0
Résidus dans les zones non admises (%)	0

**Tableau 5.18 :** Données cristallographiques globales du modèle établi pour aIF2 $\alpha\gamma$  de *S. solfataricus*. Ces statistiques proviennent de l'analyse du modèle final à l'aide du programme PROCHECK (Laskowski *et al.*, 1993).

Le modèle final ainsi construit correspond aux résidus 2 à 415 d'aIF2 $\gamma$  de *S. solfataricus* (soit la protéine entière, à l'exception de la méthionine initiatrice normalement clivée) et les résidus 1 à 264 d'aIF2 $\alpha$  de *S. solfataricus* (protéine entière moins les deux lysines C-terminales). Dans la structure d'aIF2 $\gamma$ , deux zones sont absentes : la région

comprenant les résidus 36 à 43 et celle comprenant les résidus 262 à 266. Par contre, le GDPNP, l'ion  $\text{Mg}^{2+}$  qui le stabilise ainsi que l'ion  $\text{Zn}^{2+}$  du domaine du zinc sont parfaitement visibles. Au sein d' $\alpha\text{IF}2\alpha$ , seule la région des résidus 48 à 52 n'est pas visualisée dans la densité électronique. 42 molécules d'eau ont également été placées dans le modèle structural. Les chaînes latérales de tous les résidus ont été positionnées dans la densité électronique et tous les résidus adoptent une conformation favorable comme en témoigne le diagramme de Ramachandran (**Figure 5.19**). Enfin, des facteurs d'agitation thermique de 59.4 et 40.1 Å<sup>2</sup> ont été obtenus pour  $\alpha$  et  $\gamma$  respectivement.



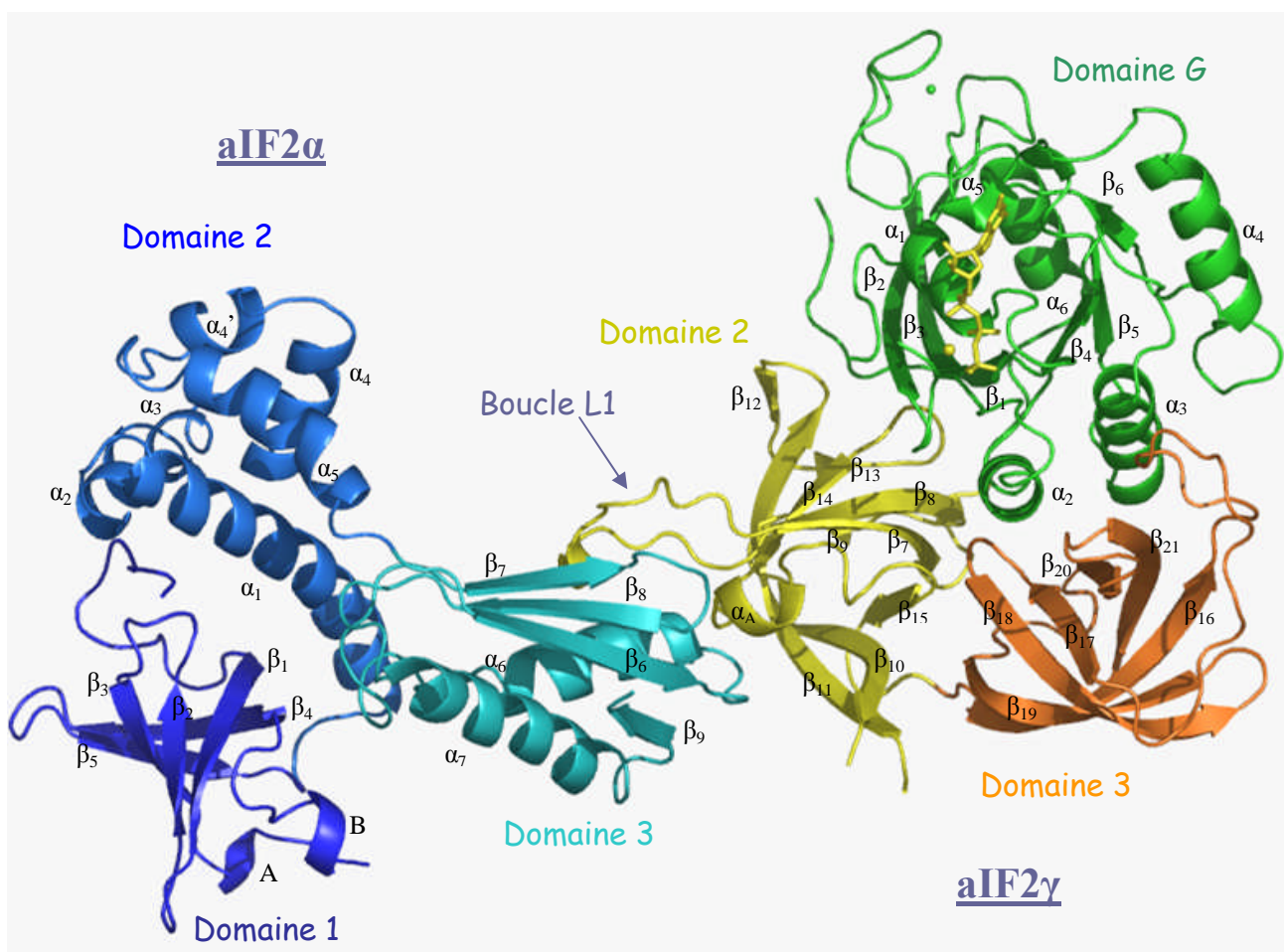
**Figure 5.19 :** Diagramme de Ramachandran du modèle d' $\alpha\text{IF}2\alpha\gamma$  de *S. solfataricus* affiné à 3.0 Å de résolution. Les zones en rouge représentent les zones favorables (76.6%), les zones en jaune vif indiquent les zones tolérées (23.4%) et les zones en jaune pâle correspondent aux régions généralement tolérées. Enfin, les régions non admises sont indiquées en blanc. Les différents résidus sont symbolisés par des carrés à l'exception des glycines représentées par un triangle. Ce diagramme a été réalisé à l'aide du programme PROCHECK (Laskowski *et al.*, 1993).

## Partie D : Structure de l'hétérodimère aIF2 $\alpha\gamma$ de l'archée *Sulfolobus solfataricus*.

### 1) Description générale.

#### 1. Organisation structurale de l'hétérodimère.

La structure de l'hétérodimère aIF2 $\alpha\gamma$  de l'archée *Sulfolobus solfataricus*, déterminée à 3.0 Å de résolution, comprend les deux sous-unités  $\alpha$  et  $\gamma$  entières (à l'exception des résidus 36 à 43, 262 à 266 pour  $\gamma$  et 48 à 52 pour  $\alpha$ ). Il s'agit d'une structure relativement étendue : les deux protéines interagissent au niveau d'une unique zone de contact. La **Figure 5.20** représente la structure tridimensionnelle du complexe.



**Figure 5.20 :** Structure de l'hétérodimère aIF2 $\alpha\gamma$  de *Sulfolobus solfataricus* à 3.0 Å de résolution. Figure dessinée avec PyMOL. Les domaines de Ss-aIF2 $\gamma$  sont représentés en vert, jaune et orange. Le GDPNP est représenté en jaune au sein du domaine G. Les domaines de Ss-aIF2 $\alpha$  sont représentés en différentes nuances de bleu.

Dans son ensemble, la sous-unité  $\gamma$  de *S. solfataricus* présente une organisation similaire à celles d'aIF2 $\gamma$  de *P. abyssi* et aIF2 $\gamma$  de *M. jannaschii* (Schmitt *et al.*, 2002 ; Roll-Mecak *et al.*, 2004). Les trois domaines structuraux (le domaine G et les deux tonneaux  $\beta$ ) sont retrouvés et leur agencement en une structure compacte, avec le domaine G à proximité des domaines II et III, est très proche de celui observé au sein des deux autres sous-unités  $\gamma$  ainsi que dans le complexe EF1A:GDPNP (Nissen *et al.*, 1995). Le domaine de fixation du Zn à l'extérieur du domaine G est également présent.

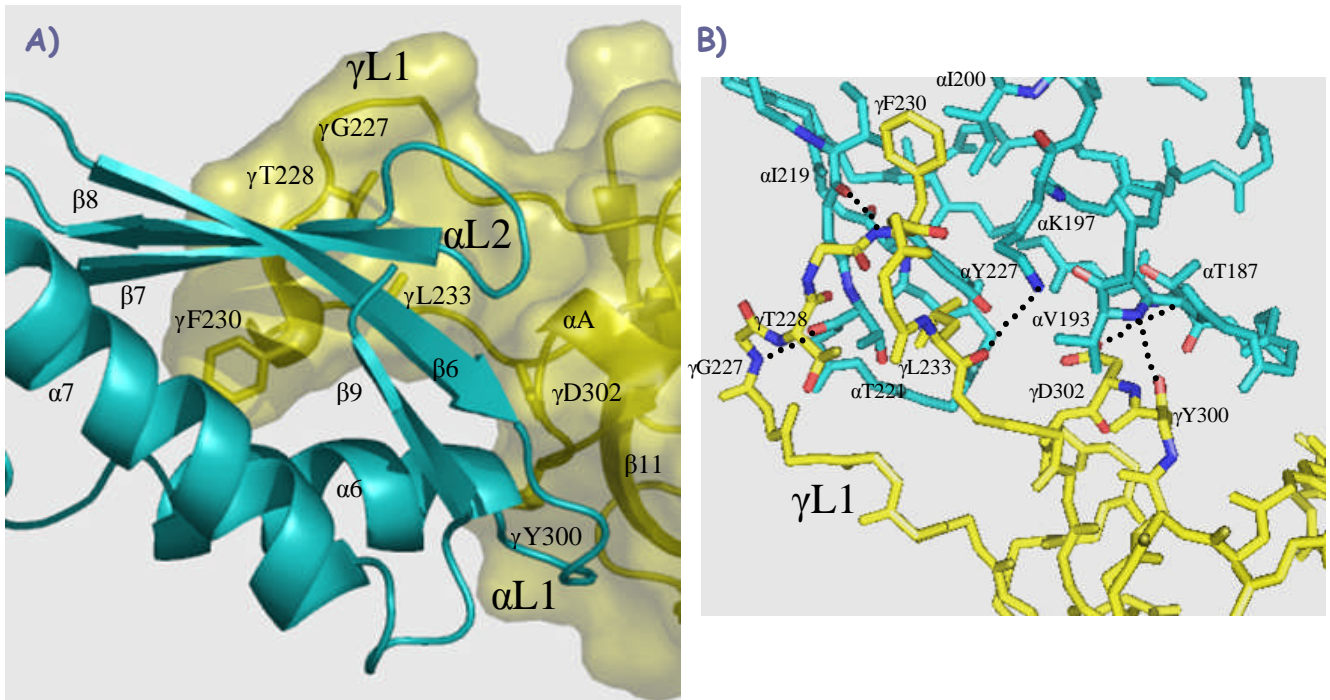
L'organisation structurale de chacun des trois domaines de la sous-unité  $\alpha$  est elle aussi similaire à celle rencontrée dans les autres structures d'e/aIF2 $\alpha$ . En particulier, chaque domaine séparé a pu être superposé au domaine correspondant au sein de la structure de la protéine aIF2 $\alpha$  de *P. abyssi* décrite au **Chapitre IV**. L'orientation relative des domaines 1 et 2 de la protéine est également conservée par rapport à celle rencontrée au sein des protéines humaine, de levure ou de *P. abyssi* (Ito *et al.*, 2004 ; Nonato *et al.*, 2002 ; Dhaliwal and Hoffman, 2003 ; et ce travail, **Chapitre IV**). Notamment, les deux résidus Lys93 et Glu33 impliqués dans une interaction stabilisante entre les deux domaines d'aIF2 $\alpha$  de *P. abyssi* sont conservés chez aIF2 $\alpha$  de *S. solfataricus* et jouent un rôle similaire de maintien de l'orientation relative des deux domaines N-terminaux de la protéine.

La superposition des structures d'aIF2 $\alpha$  de *P. abyssi* et d'aIF2 $\alpha$  de *S. solfataricus* révèle une rotation du domaine 3 d' $\alpha$  de près de 180° autour de l'angle dièdre N-C $_{\alpha 174}$ -C $_{\alpha 175}$ -N. La mobilité du bloc formé par les domaines 1 et 2 d' $\alpha$  vis-à-vis du domaine 3 est donc retrouvée au sein de Ss-aIF2 $\alpha$ . Elle pourrait avoir une importance fonctionnelle significative pour le relais d'information entre aIF2 et ses autres partenaires au cours du démarrage de la traduction comme nous l'avons déjà discuté au **Chapitre IV**.

## **2. Zone d'interaction entre les deux sous-unités.**

Au sein de l'hétérodimère Ss-aIF2 $\alpha\gamma$ , les deux sous-unités  $\alpha$  et  $\gamma$  n'interagissent qu'au niveau d'une seule zone de contact. Les régions impliquées dans cette liaison correspondent au domaine 2 d'aIF2 $\gamma$  et au domaine 3 d'aIF2 $\alpha$ . La surface de contact représente environ 1700 Å<sup>2</sup>. La **Figure 5.21** présente une vue rapprochée de la zone de contact.

Deux boucles du domaine 3 de Ss-aIF2 $\alpha$  sont ancrées dans deux poches formées à la surface du domaine 2 de Ss-aIF2 $\gamma$  : il s'agit des boucles  $\alpha$ L1 (reliant le brin  $\beta$ 6 à l'hélice  $\alpha$ 6) et  $\alpha$ L2 (entre les brins  $\beta$ 7 et  $\beta$ 8). Au sein du domaine 2 de Ss-aIF2 $\gamma$ , la première poche de fixation est délimitée par la boucle  $\gamma$ L1 (résidus 224 à 234), la seconde par le brin  $\beta$ 11. La petite hélice  $\alpha$ A de Ss-aIF2 $\gamma$  se situe à la charnière des deux poches de fixation.



**Figure 5.21 :** Vue rapprochée de la zone d'interaction entre les sous-unités  $\alpha$  et  $\gamma$  au sein du dimère  $\alpha\text{IF}2\alpha\gamma$  de *S. solfataricus*. Ss- $\alpha\text{IF}2\alpha$  est représentée en cyan, Ss- $\alpha\text{IF}2\gamma$  en jaune. **A)** Vue générale : mise en évidence des éléments de structure impliqués dans l'interaction. **B)** Interactions électrostatiques entre  $\alpha$  et  $\gamma$ .

L'interaction entre les deux sous-unités implique donc la grande boucle L1 d' $\alpha\text{IF}2\gamma$ . En particulier, des liaisons H sont formées entre les groupes NH des chaînes principales des résidus  $\gamma\text{T}228$  et  $\gamma\text{G}227$  et le groupe CO de la chaîne principale du résidu  $\alpha\text{T}221$ . De même,  $\gamma\text{F}230$  interagit avec  $\alpha\text{I}219$ . De plus, la chaîne latérale de  $\gamma\text{F}230$  est accolée à  $\alpha\text{I}200$ . La chaîne latérale de  $\gamma\text{L}233$  est impliquée dans des contacts de Van der Waals avec les chaînes latérales de  $\alpha\text{T}221$  et du résidu strictement conservé  $\alpha\text{Y}227$ . Le groupe amide de la chaîne principale de  $\gamma\text{L}233$  interagit également avec l'atome d'azote N $\epsilon$  de  $\alpha\text{K}197$ . Dans toutes les séquences d' $e/\alpha\text{IF}2\alpha$ , un résidu Lys ou un résidu Arg est d'ailleurs rencontré au niveau de la position équivalente à K197.

Au sein de la seconde poche de liaison, le groupe carbonyle de  $\gamma\text{Y}300$  interagit avec le groupe NH de  $\alpha\text{V}193$  et la chaîne latérale de  $\gamma\text{D}302$  interagit avec le groupe amide de  $\alpha\text{T}187$ . Cette seconde région implique également des résidus très conservés parmi tous les  $e/\alpha\text{IF}2\gamma$ . En particulier, le caractère acide du résidu occupant la position de  $\gamma\text{D}302$  est retrouvé dans toutes ces séquences (voir **Annexe 1**).

La boucle  $\gamma\text{L}1$ , spécifique des facteurs  $e/\alpha\text{IF}2\gamma$  et absente de leurs homologues de type facteur d'élongation tels qu' $\text{EF}1\text{A}$ , a déjà été identifiée dans le **Chapitre III** comme responsable de la liaison à la sous-unité  $\alpha$ . En effet, le mutant  $\gamma_{\Delta\text{L}1}$  privé de cette boucle était

incapable de s'associer *in vitro* à aIF2 $\alpha$ . De plus, des expériences de mutagenèse dirigée réalisée sur la protéine  $\gamma$  de *M. jannaschii* ont corroboré l'identification de la zone d'interaction. Ainsi les mutants  $\gamma$ L256D et  $\gamma$ D325A de *M. jannaschii* présentaient un sévère défaut d'association à  $\alpha$  (Roll-Mecak *et al.*, 2004). Ces résidus correspondent respectivement à  $\gamma$ L233 et  $\gamma$ D302 chez *S. solfataricus*. Ainsi, l'étude structurale présentée ici est en parfait accord avec les données biochimiques résultant de nos travaux sur les protéines de *P. abyssi* et de ceux menés sur aIF2 $\gamma$  de *M. jannaschii*.

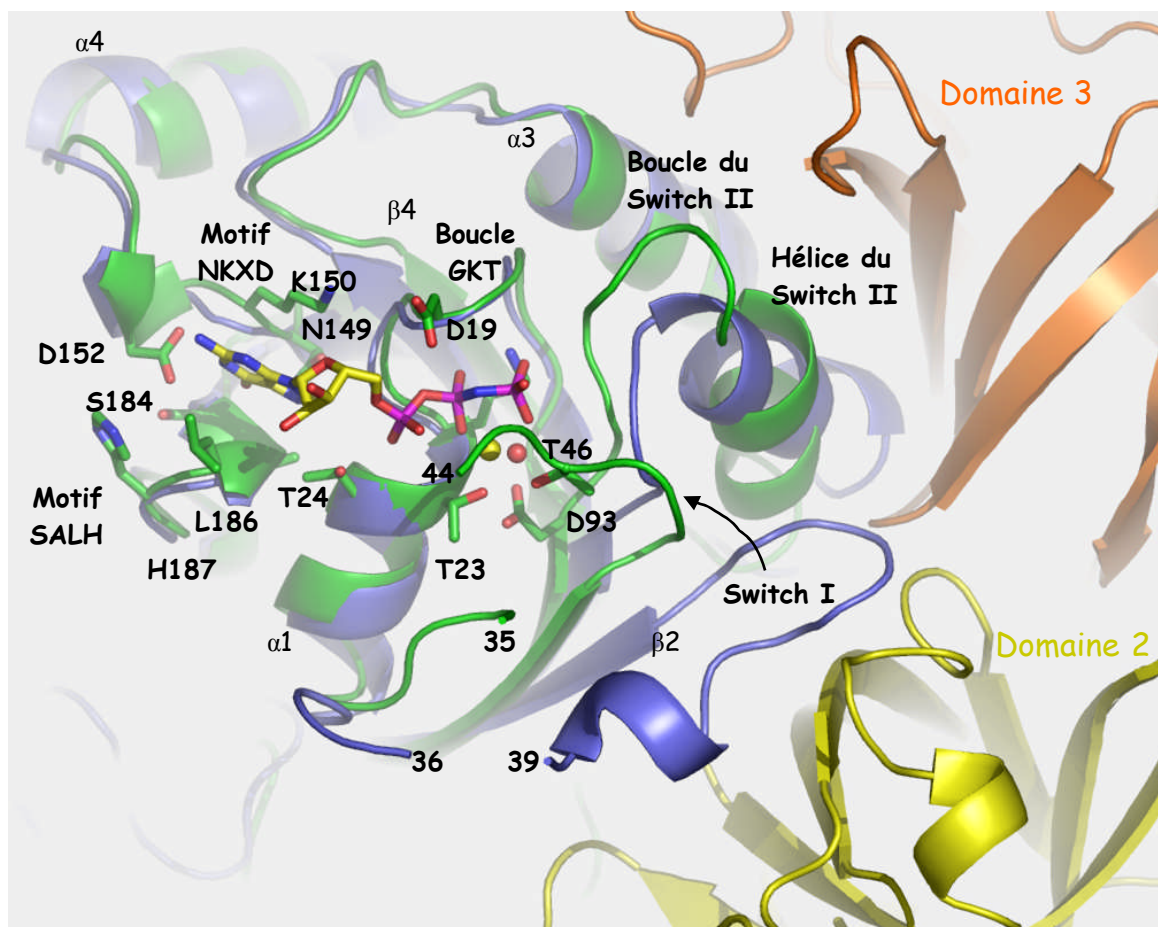
Dans la structure d'aIF2 $\gamma$  de *P. abyssi* (Schmitt *et al.*, 2002), la boucle L1 n'avait pas pu être visualisée du fait d'un désordre partiel des résidus 224 à 227. Cette boucle est par contre visible dans la structure de la protéine de *M. jannaschii* (Roll-Mecak *et al.*, 2004). La conformation adoptée par  $\gamma$ L1 de Ss-aIF2 a alors été comparée à celle observée pour la protéine de *M. jannaschii* : les deux conformations sont identiques. De plus, la comparaison des domaines 2 de chacune des trois structures d'aIF2 $\gamma$  indique que la conformation des chaînes principales est constante au sein des trois protéines : les domaines 2 et 3 d'aIF2 $\gamma$  de *P. abyssi* et de *M. jannaschii* ont été superposés aux domaines correspondants d'aIF2 $\gamma$  de *S. solfataricus* avec dans chaque cas un rmsd de 1.17Å pour 177 carbones alpha comparés. De même, le domaine 3 d'aIF2 $\alpha$  adopte une conformation similaire au sein de la protéine libre de *P. abyssi* et au sein du complexe  $\alpha\gamma$  de *S. solfataricus* (rmsd de 1.5 Å pour 83 atomes comparés entre les deux domaines 3 d' $\alpha$ ). Ces résultats suggèrent donc que la poche de liaison à  $\alpha$  est déjà formée au sein de la sous-unité  $\gamma$  isolée.

### **3. Poche de fixation des nucléotides guanine.**

La cristallisation de l'hétérodimère Ss-aIF2 $\alpha\gamma$  a été réalisée en présence de GDPNP et de magnésium. La poche de fixation de l'analogue du GTP a été identifiée sans ambiguïté (**Figure 5.22**). Les séquences consensus des protéines G sont retrouvées, en particulier les motifs  $_{184}$ SALH $_{188}$  et  $_{149}$ NKVD $_{152}$  qui entourent la base G du nucléotide et la boucle  $_{21}$ GKT $_{23}$  qui interagit avec les phosphates  $\beta$  et  $\gamma$  du nucléotide. Les deux régions Switch I et II sont également identifiables, bien qu'une partie du motif Switch I (résidus 36 à 43) ne soit pas visible dans la densité électronique.

Dans cette structure, l'ion Mg $^{2+}$  est coordonné par 5 ligands oxygènes selon une géométrie inhabituelle de type pyramidal à base carrée. Dans le plan équatorial, le magnésium est coordonné par deux oxygènes provenant des groupes phosphates  $\beta$  et  $\gamma$  du GDPNP. Deux autres ligands oxygènes sont amenés par les chaînes latérales des deux thréonines Thr23 (motif GKT) et Thr46 (Switch I). En position apicale, c'est une molécule d'eau qui vient jouer le rôle du dernier ligand. Cette molécule d'eau est maintenue en position par une

interaction avec la chaîne latérale du résidu Asp93 et par une autre interaction avec le groupe carbonyle de l'Ala94 (motif  $_{93}\text{DAPG}_{96}$  de la boucle Switch II). Dans la structure d'EF1A:GDPNP de *T. thermophilus*, une seconde molécule d'eau interagit avec l'ion magnésium qui adopte alors une coordination octaédrique classique (Nissen *et al.*, 1995). Dans le cas d'aIF2 $\gamma$  de *S. solfataricus*, cette seconde molécule d'eau n'est pas visible dans la densité électronique.



**Figure 5.22 :** Poche de fixation des nucléotides guanine dans le domaine G d'aIF2 $\gamma$  : superposition de la structure du domaine G d'aIF2 $\gamma$ :GDP de *P. abyssi* et de celle du domaine G d'aIF2 $\gamma$ :GDPNP de *S. solfataricus*. Les domaines de Ss-aIF2 $\gamma$  ont été colorés comme suit : domaine G en vert, domaine 2 en jaune, domaine 3 en orange. Pour la structure Pa-aIF2 $\gamma$ :GDP, seul le domaine G a été représenté en bleu. La molécule de GDPNP représentée est celle de la structure de Ss-aIF2 $\gamma$ :GDPNP. L'ion Mg<sup>2+</sup> est représenté par une sphère jaune et la molécule d'eau qui le ligande est représentée par une sphère rouge. Les motifs caractéristiques des protéines G impliqués dans l'interaction avec le GDPNP et l'ion Mg<sup>2+</sup> sont indiqués. Le mouvement des régions Switch entre l'état GDP et l'état GDPNP est indiqué par une flèche. La numérotation adoptée est celle de *S. solfataricus*. Figure réalisée avec PyMol.

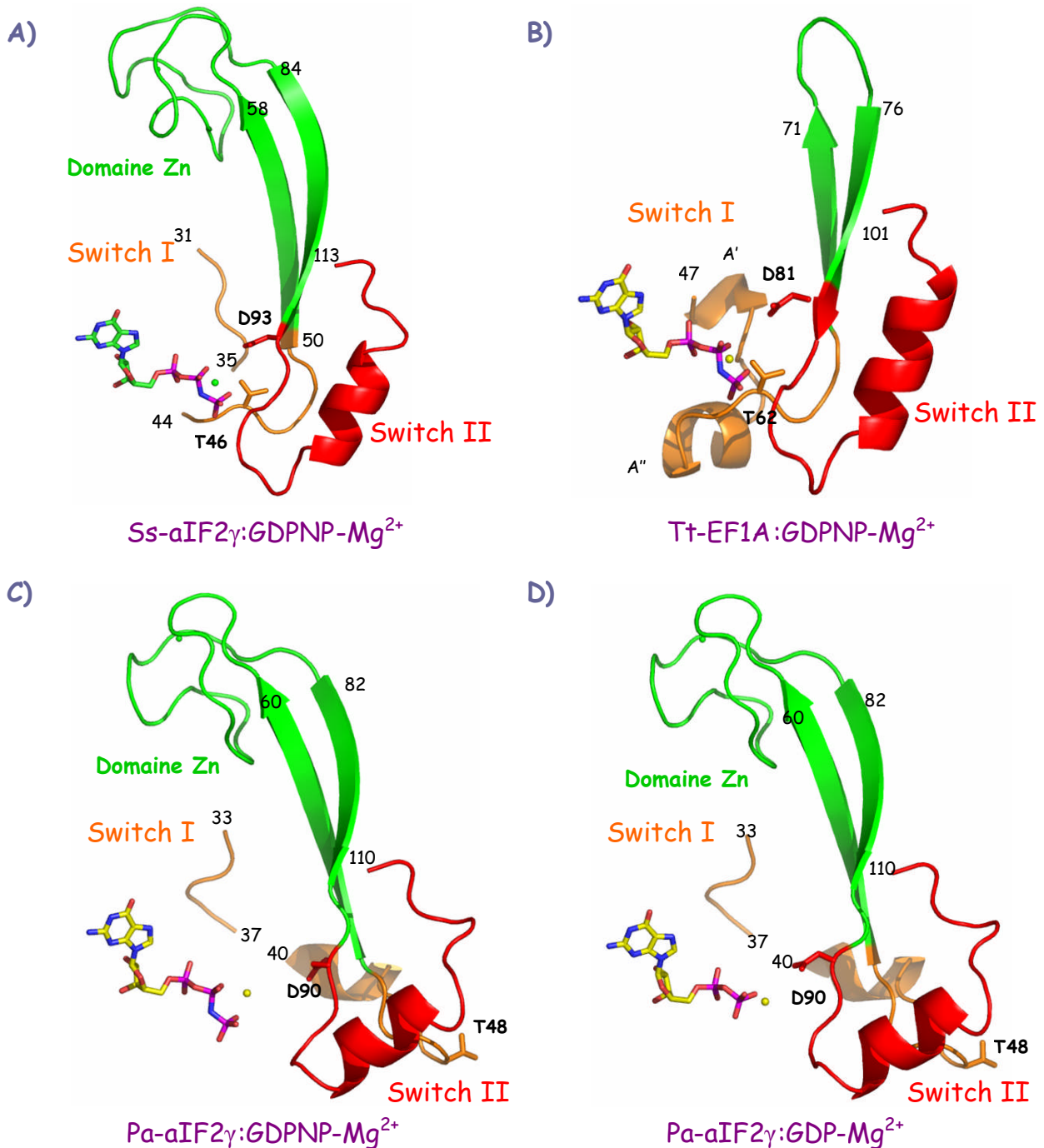
#### 4. Les régions Switch I et Switch II.

La conformation des régions Switch au sein de la structure de Ss-aIF2 $\gamma$ :GDPNP a été comparée à celle des mêmes régions chez Pa-aIF2 $\gamma$ :GDP (**Figure 5.22**). Pour la première fois, un mouvement significatif des régions Switch est observé entre les deux états ON et OFF de la protéine. Lors du passage de la forme GDP à la forme GTP, l'hélice de la région Switch II ( $\alpha 2$ ) subit un déplacement non négligeable. En effet, chez *S. solfataricus*, cette hélice est décalée de 3 résidus au sein de la séquence en acides aminés de  $\gamma$  par rapport à la protéine de *P. abyssi* ( $\alpha 2$  de *S. solfataricus* : <sub>101</sub>MATMLSGA<sub>108</sub> ;  $\alpha 2$  de *P. abyssi* : <sub>94</sub>EALMTTML<sub>101</sub>). Chez *S. solfataricus*, comparativement à *P. abyssi*, l'extrémité N-terminale d' $\alpha 2$  est déroulée, ce qui permet d'étendre la boucle précédant  $\alpha 2$  et par conséquent, l'interaction entre l'ion magnésium et l'Asp93 d'une part, et l'Ala94 d'autre part, peut avoir lieu. De la même manière, le résidu Gly96 interagit désormais avec le phosphate  $\gamma$  du GDPNP. Une autre transition, de la région Switch I cette fois-ci, permet de rapprocher la boucle Switch I du nucléotide guanine et autorise ainsi l'interaction entre Thr46 et l'ion Mg<sup>2+</sup>.

La comparaison de la position qu'adoptent ces régions Switch avec celle rencontrée chez aIF2 $\gamma$  de *P. abyssi*, chez aIF2 $\gamma$  de *M. jannaschii* ou chez les facteurs d'élongations EF1A montre que la conformation de ces deux régions Switch au sein de Ss-aIF2 $\gamma$ :GDPNP de *S. solfataricus* est similaire à celle décrite pour les complexes EF1A:GDPNP et EF1A:GDPNP:ARNt (Nissen *et al.*, 1995 ; Kjeldgaard *et al.*, 1993 ; Berchtold *et al.*, 1993) (**Figure 5.24**). En particulier, l'hélice  $\alpha 2$  du Switch 2 est identiquement positionnée chez aIF2 $\gamma$  de *S. solfataricus* et chez EF1A:GDPNP alors que son orientation est modifiée dans les deux structures d'aIF2 $\gamma$  de *P. abyssi*. L'hélice A'' rencontrée chez EF1A n'est pas visible dans la structure d'aIF2 $\gamma$  de *S. solfataricus* (résidus 36 à 43 non construits car non visibles dans la densité électronique). Chez EF1A, cette hélice fait partie de la région effectrice qui interagit avec l'ARNt mais également avec l'ion Mg<sup>2+</sup> par l'intermédiaire d'une seconde molécule d'eau. Cette seconde molécule d'eau forme en effet une liaison H avec le résidu Asp91 chez EF1A de *T. thermophilus* (Nissen *et al.*, 1995). Chez aIF2 $\gamma$  de *S. solfataricus*, ce second ligand H<sub>2</sub>O n'est pas visualisé dans la densité électronique. De ce fait, la coordination octaédrique préférentielle de l'ion magnésium observée dans les complexes EF1A:GDPNP (Nissen *et al.*, 1995 ; Kjeldgaard *et al.*, 1993 ; Berchtold *et al.*, 1993) apparaît ici comme une géométrie de type pyramidale à base carrée.

Chez EF1A, les deux hélices A' et A'' participent à l'interaction avec le polynucléotide ARNt. L'absence d'ARNt dans la structure du complexe Ss-aIF2 $\alpha\gamma$  pourrait expliquer pourquoi l'hélice A'' de Ss-aIF2 $\gamma$  ne peut adopter une conformation stable. En présence du Met-ARNt<sub>i</sub><sup>Met</sup>, A'' pourrait alors se stabiliser et le résidu Ser36 (qui occupe la

position équivalente à D51 d'EF1A dans la séquence d'aIF2 $\gamma$ ) pourrait alors venir stabiliser l'ion magnésium en rétablissant son hexa-coordination.



**Figure 5.23 :** Conformation des régions Switch au sein d'aIF2 $\gamma$  et d'EF1A. Les différentes représentations sont déduites de la superposition des domaines G d'aIF2 $\gamma$  de *S. solfataricus*, d'EF1A au sein du complexe EF1A:GDPNP:ARNt (Nissen *et al.*, 1995) et d'aIF2 $\gamma$  de *P. abyssi* sous forme GDPNP et sous forme GDP (Schmitt *et al.*, 2002). Dans chaque structure, la région Switch I est colorée en orange et la région Switch II en rouge. **A)** aIF2 $\gamma$ :GDPNP de *S. solfataricus* : Switch I (G31 à G50) et Switch II (D93 à G113). **B)** EF1A:GDPNP (Code PDB : 1TTT) : Switch I (Y47 à A66) et Switch II (D81 à G101). **C)** aIF2 $\gamma$ :GDPNP de *P. abyssi* : Switch I (G33 à G50) et Switch II (D90 à G110). **D)** aIF2 $\gamma$ :GDP de *P. abyssi* : Switch I (G33 à G50) et Switch II (D90 à G110). Seul un très faible réarrangement des régions Switch est observé entre **C)** et **D)**. La région la plus affectée est la boucle du Switch II qui, après fixation du GDPNP, se déplace de telle manière que D90 n'interagit plus avec l'ion magnésium.

## 2) Positionnement de l'ARNt au sein du complexe Ss-aIF2 $\alpha\gamma$ .

### 1. Modèle de positionnement.

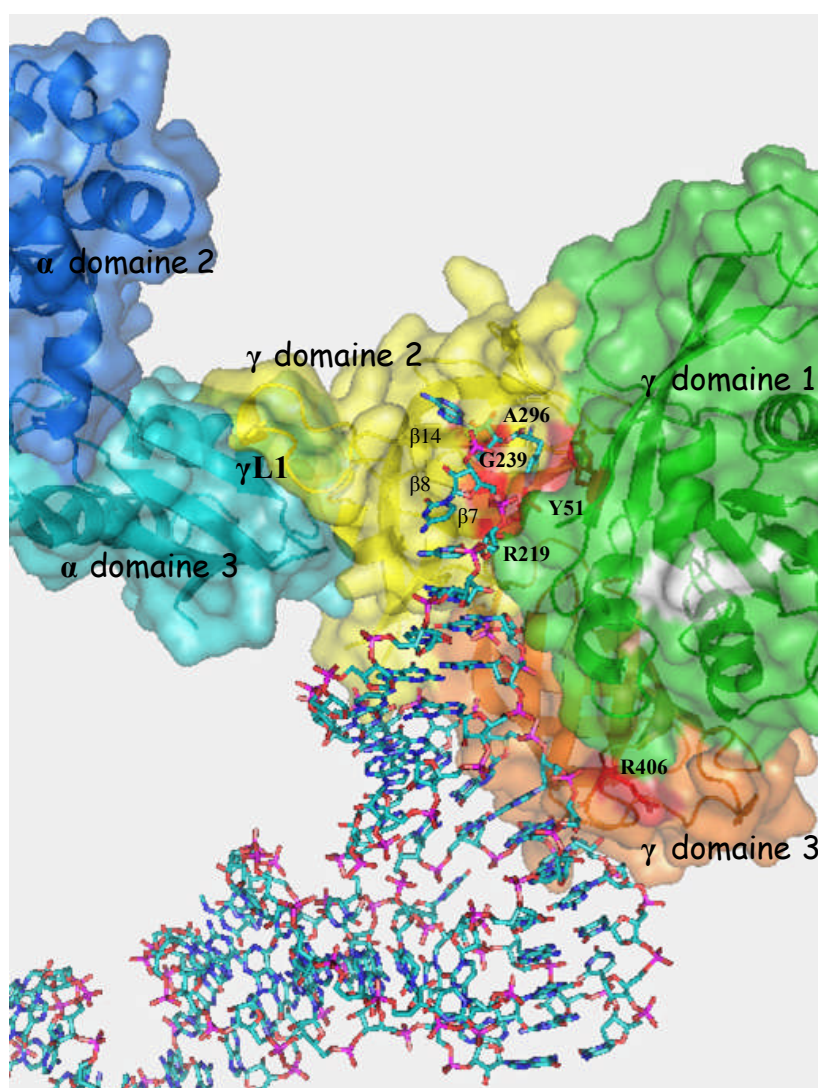
L'étude des mouvements des Switch d'aIF2 $\gamma$  a montré qu'au sein de l'hétérodimère Ss-aIF2 $\alpha\gamma$ :GDPNP,  $\gamma$  adoptait une conformation très proche de celle d'EF1A complexé au GDPNP et à un ARNt élongateur. C'est d'ailleurs au sein de cette structure  $\alpha\gamma$  que la plus grande similarité est rencontrée avec EF1A en conformation active. Un rmsd de 1.64 Å pour 302 C $\alpha$  comparés est obtenu lors de la superposition d'aIF2 $\gamma$  de *S. solfataricus* avec EF1A:GDPNP. Un rmsd de seulement 1.95 Å pour 253 atomes comparés est obtenu si Ss-aIF2 $\gamma$  est remplacé par Pa-aIF2 $\gamma$  ou Mj-aIF2 $\gamma$ .

En réalité, dans la structure de l'hétérodimère Ss-aIF2 $\alpha\gamma$ , le mouvement de la région Switch I permet l'ouverture d'un tunnel dans lequel la tige acceptrice 3' de l'ARNt peut s'engager pour atteindre la poche de fixation du résidu méthionine dans le domaine 2 d'aIF2 $\gamma$  (poche de fixation identifiée par homologie avec EF1A et confirmée par mutagenèse dirigée dans le **Chapitre III** de ce manuscrit). Un modèle de positionnement encore mieux ajusté de l'ARNt sur ce complexe Ss-aIF2 $\alpha\gamma$  a donc été directement déduit de la superposition de la structure de l'hétérodimère Ss-aIF2 $\alpha\gamma$  à la structure complète du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995) (**Figure 5.24**).

Le modèle de positionnement ainsi construit n'interfère à aucun endroit avec la structure de Ss-aIF2 $\gamma$  et/ou Ss-aIF2 $\alpha$  et aucun ajustement par rapport au modèle déduit de la seule superposition de  $\gamma$  avec EF1A n'a été nécessaire. Notamment, aucune réorientation des domaines 2 et 3 de  $\gamma$  par rapport au domaine 1 n'a été nécessaire pour permettre l'insertion de la tige acceptrice de l'ARNt au sein de la protéine. La **Figure 5.24** montre que la base terminale A76 de l'ARNt vient se fixer dans la poche du domaine 2 de  $\gamma$  formée par les brins  $\beta$ 11 et  $\beta$ 14. L'acide aminé chargé sur l'ARNt vient quant à lui se loger dans un creux formé par les brins  $\beta$ 7,  $\beta$ 8 et  $\beta$ 14 d'un côté et par la région Switch I de l'autre côté.

Ce modèle de positionnement permet de plus d'expliquer les résultats biochimiques obtenus avec divers mutants de la sous-unité  $\gamma$ . Ainsi, les mutants suivants présentaient un sévère défaut d'interaction avec la molécule d'ARNt : Y142H chez la levure (Erickson *et al.*, 1997), R319D et G397A chez *M. jannaschii* (Roll-Mecak *et al.*, 2004), G235D chez *P. abyssi* (Schmitt *et al.*, 2002 et cette étude) et R510H chez la levure (Dorris *et al.*, 1995). Ces résidus correspondent respectivement à Y51, R219, A296, G239 et R406 chez Ss-aIF2 $\gamma$  (voir **Figure 5.24** pour leur localisation). Dans la structure du complexe Ss-aIF2 $\alpha\gamma$ , Y51 appartient à la région Switch I et, d'après le modèle de positionnement, il serait accolé à l'acide aminé porté

par l'ARNt. Les résidus R219 ( $\beta$ 7), G239 ( $\beta$ 8) et A296 ( $\beta$ 14) contribueraient à la liaison de l'acide aminé et de la base terminale A76. R406 est situé assez loin de cette zone, dans le domaine 3 d'aIF2 $\gamma$ . Cependant, d'après notre modèle, ce résidu interagirait avec le bras T de l'ARNt. Un fait remarquable dans la structure d'aIF2 $\gamma$  de *S. solfataricus* est que les chaînes latérales des résidus Y51, R219 (Switch I) et également F221 (domaine 2) adoptent une position totalement différente de celle qu'ils occupent dans les structures d'aIF2 $\gamma$  de *P. abyssi* ou de *M. jannaschii*. Ces nouvelles positions pourraient résulter des mouvements des deux régions Switch observés dans la structure de l'hétérodimère  $\alpha\gamma$  et ils pourraient refléter une réorganisation du site de fixation de l'ARNt au sein du dimère aIF2 $\alpha\gamma$ .



**Figure 5.24 :** Modèle de positionnement du Phe-ARNt<sup>Phe</sup> sur la structure d'aIF2 $\alpha\gamma$ :GDPNP de *S. solfataricus* à partir de la structure du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995). La surface moléculaire du dimère  $\alpha\gamma$  est représentée en transparence. Les domaines de chaque sous-unité sont colorés suivant le même code que dans la **Figure 5.20**. La surface correspondant à la molécule de GDPNP est représentée en blanc au sein du domaine G de  $\gamma$ . La boucle L1 et les résidus importants pour la reconnaissance de l'ARNt sont indiqués.

## 2. Rôle de la sous-unité $\alpha$ dans la liaison de l'ARNt.

Comme pour *P. abyssi*, le domaine 3 d'aIF2 $\alpha$  de *S. solfataricus* est nécessaire pour conférer à Ss-aIF2 $\gamma$  une affinité optimale vis-à-vis de l'ARNt initiateur. Deux hypothèses sont envisageables pour expliquer le rôle joué par  $\alpha_{D3}$  dans le renforcement de cette affinité : le domaine 3 d'aIF2 $\alpha$  pourrait interagir directement avec le polynucléotide ou bien son action pourrait être indirecte, c'est-à-dire qu' $\alpha$  pourrait induire une réorganisation conformationnelle de  $\gamma$  qui serait alors plus affine pour l'ARNt. Les expériences de retard de migration de l'ARNt sur gel de polyacrylamide décrites au **Chapitre III** ont montré que le domaine 3 d'aIF2 $\alpha$  présente une capacité de liaison des ARNs qui est en faveur d'un contact direct entre  $\alpha_{D3}$  et le polynucléotide. Cependant, cette propriété n'est pas spécifique du motif ARNt. De plus, des expériences de titrage par RMN en présence du domaine 3 d'eIF2 $\alpha$  humain marqué avec de l'azote  $^{15}\text{N}$  et en présence d'ARNt $_{i}^{\text{Met}}$  libre n'ont révélé aucune interaction entre les deux molécules (Ito *et al.*, 2004). Au vu de ces résultats, le mode de participation d' $\alpha$  restait donc assez flou et nécessitait une étude plus approfondie.

La structure de l'hétérodimère Ss-aIF2 $\alpha\gamma$  décrite dans ce chapitre ainsi que le modèle de positionnement de l'ARNt réalisé sur le complexe ne nous permettent pas d'envisager une interaction directe entre les deux molécules. Comme le montre la **Figure 5.24**, dans cette conformation, le domaine 3 d' $\alpha$  ne présente aucun résidu suffisamment proche pour être engagés dans une interaction directe avec le polynucléotide. Deux hypothèses restent toutefois plausibles : lors de la fixation de la molécule d'ARNt sur aIF2 $\gamma$ , un changement conformationnel pourrait avoir lieu et permettre le rapprochement d' $\alpha$  de l'ARNt, autorisant ainsi un contact direct entre les deux partenaires. Ou bien, encore une fois, la participation d' $\alpha$  pourrait être indirecte.

Pour tester la première idée, nous avons cherché sur la surface d'aIF2 $\alpha_{D3}$  les résidus pointant vers la molécule d'ARNt et qui, en cas de changement conformationnel, pourraient se retrouver en contact direct avec le polynucléotide. Deux résidus apparaissent ainsi être de bons candidats pour une interaction potentielle : il s'agit des résidus Arg186 et Arg226. Ces deux résidus chargés positivement pourraient former des interactions électrostatiques avec le squelette phosphaté de l'ARNt en cas d'un rapprochement du domaine 3 d'aIF2 $\alpha$ . Le résidu R226 est d'ailleurs très conservé parmi les différents e/aIF2 $\alpha$  et un résidu de même polarité est généralement rencontré en son absence (Lys, Asn).

En conséquence, les deux mutants ponctuels  $\alpha_{R186A}$  et  $\alpha_{R226A}$  ont été construits par mutagenèse dirigée par la méthode *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene à partir du vecteur cible pET3a $\alpha$ sslpa. Les protéines résultantes ont été purifiées selon le même

protocole que la protéine native. Dans un premier temps, leur capacité d'assemblage avec la sous-unité  $\gamma$  a été testée sur gel de polyacrylamide natif. Après vérification de l'appariement entre les deux sous-unités pour chaque mutant, l'hétérodimère correspondant a été purifié. Puis, les hétérodimères mutés ont été testés pour leur capacité à lier l'ARNt initiateur. Le test de protection contre la déacylation a été reproduit en présence de ces différents mutants à 51°C et à 2 nM en  $^{35}\text{S}$ -Met-ARNt<sub>f</sub><sup>Met</sup>. Le **Tableau 5.25** reprend les valeurs de Kd obtenues.

	$\alpha$	$\beta$	$\gamma$	Kd (nM)
1	×	×	×	1.5 ± 0.5
2	×		×	7.5 ± 1.0
3	$\alpha_{\text{R186A}}$		×	7.4 ± 1.0
4	$\alpha_{\text{R226A}}$		×	27 ± 3.5
5	$\alpha_{\text{Y227A}}$		×	23 ± 1.5

**Tableau 5.25 :** Constantes de dissociation des assemblages d'aIF2 de *S. solfataricus* mentionnés en présence du Met-ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* 2 nM et de 1 mM GDPNP-Mg<sup>2+</sup>. Les sous-unités présentes dans l'assemblage testé, natives ou mutées, sont indiquées par une croix.

L'hétérodimère Ss-aIF2 $\alpha_{\text{R186A}}\gamma$  protège l'ARNt initiateur de la déacylation avec la même efficacité que l'hétérodimère natif. Par contre, l'affinité de Ss-aIF2 $\alpha_{\text{R226A}}\gamma$  pour le Met-ARNt<sub>f</sub><sup>Met</sup>, quoique relativement élevée, est réduite d'un facteur 3 par rapport à celle du complexe natif. Cependant, cette variation, bien que significative, reste trop faible pour permettre d'envisager une liaison directe d' $\alpha_{\text{D3}}$  à l'ARNt par l'intermédiaire du résidu Arg226. Il faut également noter qu'un réarrangement conformationnel permettant une interaction directe entre aIF2 $\alpha$  et l'ARNt nécessiterait une distorsion importante de l'interface  $\alpha/\gamma$ . Une telle distorsion semble peu probable du fait de la rigidité apparente de cette interface.

### 3) Les sous-unités $\alpha$ et $\gamma$ : un rôle concerté au sein de l'hétérotrimère ?

Jusqu'à présent, seules deux structures d'aIF2 $\gamma$  étaient disponibles dans la littérature : celle d'aIF2 $\gamma$  de *P. abyssi*, sous forme libre ou complexée à du GDPNP ou du GDP (Schmitt *et al.*, 2002) et celle d'aIF2 $\gamma$  de *M. jannaschii* sous forme libre (Roll-Mecak *et al.*, 2004). Dans toutes ces structures, une conformation fermée des trois domaines est observée et cette conformation est similaire à celle rencontrée pour EF1A au sein du complexe actif EF1A:GDPNP (Nissen *et al.*, 1995). Cette conformation étant retrouvée dans les différentes structures d'aIF2 $\gamma$ , il est peu probable qu'elle soit due à un effet d'empilement au sein des cristaux de chaque protéine. Ainsi, la conformation ouverte observée chez EF1A:GDP ne semble pas exister pour la sous-unité  $\gamma$  du facteur de démarrage.

Dans le cas, d'aIF2 $\gamma$  de *P. abyssi*, de très faibles changements conformationnels sont observés entre les formes GDP et GDPNP (**Figure 5.23.C et D**). Les régions Switch, où se localisent généralement les réorganisations les plus significatives, ne sont pratiquement pas modifiées lors de cette transition. La structure d'aIF2 $\gamma$  de *S. solfataricus* obtenue au sein de l'hétérodimère  $\alpha\gamma$  constitue le premier exemple pour lequel les régions Switch adoptent une conformation différente et cette nouvelle conformation est identique à la conformation des Switch observée au sein du complexe EF1A:GDPNP. Dans le cas de *P. abyssi* ou de *M. jannaschii*, bien que la conformation globale de la protéine ressemble à celle d'EF1A:GDPNP, les régions Switch adoptaient un repliement plus proche de l'état OFF des protéines G, et par conséquent, moins affin pour l'ARNt. Il faut noter que dans le cas de *P. abyssi*, les structures en présence de GDP et de GDPNP ont été obtenues par trempage de cristaux libres de tous nucléotides dans des solutions contenant ces nucléotides guanine, et également par ensemencement à partir de cristaux natifs. Les changements conformationnels normalement induits par la fixation du GTP/GDP pourraient ainsi avoir été restreints du fait de l'empilement cristallin déjà formé. Cette hypothèse ne peut être exclue. Cependant, une autre explication est à envisager : au sein de l'hétérodimère  $\alpha\gamma$ , la présence d'aIF2 $\alpha$  pourrait aider la sous-unité  $\gamma$  à atteindre une conformation stable des régions Switch induite par la fixation du GDPNP et correspondant à un état ON de la protéine. Cette hypothèse expliquerait un rôle indirect d'aIF2 $\alpha$  dans la liaison de l'ARNt.

Une dernière mutation a été étudiée au sein de la sous-unité Ss-aIF2 $\alpha$  : Y227A. Le mutant  $\alpha_{Y227A}$  a été construit de la même manière que les deux autres mutants ponctuels  $\alpha_{R186A}$  et  $\alpha_{R226A}$ . Ce résidu Y227 se trouve au niveau de la zone d'interaction entre Ss-aIF2 $\alpha$  et Ss-aIF2 $\gamma$ , au sein de la boucle  $\alpha$ L2. Il établit notamment une interaction de type Van der Waals avec le résidu Leu233 de Ss-aIF2 $\gamma$ . Sa mutation en alanine n'entraîne pas de défaut d'association au sein de l'hétérodimère  $\alpha\gamma$  (confirmation par assemblage sur gel natif acide). Par contre, la mutation réduit l'affinité de l'hétérodimère  $\alpha\gamma$  muté pour le Met-ARNt<sub>f</sub><sup>Met</sup> d'un facteur 3 (**Tableau 5.25**). Ainsi, une perturbation même légère au niveau de l'interface entre aIF2 $\alpha$  et aIF2 $\gamma$  est suffisante pour moduler l'affinité du complexe pour l'ARNt initiateur. Dans le même ordre d'idée, il faut rapprocher le comportement de ce mutant de celui du mutant  $\gamma_{\Delta LI}$  obtenu pour aIF2 $\gamma$  de *P. abyssi* et décrit au **Chapitre III**. Ce mutant, qui présente un défaut d'association vis-à-vis d'aIF2 $\alpha$ , est également dans l'incapacité d'interagir avec le Met-ARNt<sub>f</sub><sup>Met</sup>. Ce résultat suggère donc un lien étroit entre l'intégrité du site de liaison d'aIF2 $\alpha$  sur  $\gamma$  et l'acquisition par aIF2 $\gamma$  d'une conformation productive optimale pour la liaison de l'ARNt.

L'absence de réarrangements conformationnels majeurs au sein d'aIF2 $\gamma$  selon la nature du nucléotide guanine lié est relativement surprenante par rapport au cas d'EF1A. Pourtant, la GTP-dépendance de la fixation de l'ARNt est confirmée. En effet, chez la levure, l'affinité d'eIF2 pour le Met-ARNt<sup>Met</sup> est modulée d'un facteur 20 selon que le nucléotide lié est le GTP ou le GDP. L'état pour lequel  $\gamma$  est complexé au GTP pourrait d'ailleurs favoriser des interactions avec le résidu méthionine, non réalisées avec la forme GDP d'eIF2 $\gamma$ , tandis que le site de fixation du corps du polynucléotide resterait identique entre les deux états de la protéine (Kapp and Lorsch, 2004). L'importance décisive du résidu méthionine dans la reconnaissance spécifique de l'ARNt initiateur serait ainsi corrélée à la nécessité pour  $\gamma$  d'être sous forme GTP. Dans le cas de l'hétérotrimère aIF2 de *S. solfataricus*, une comparaison des affinités du facteur pour l'ARNt en présence de GDPNP et en présence de GDP a également été réalisée (**Tableau 5.26**).

	$\alpha$	$\beta$	$\gamma$	Kd (nM)
<b>GDPNP</b>	×		×	7.5 ± 1.0
<b>GDP</b>	×		×	600 ± 20

**Tableau 5.26 :** Comparaison de l'affinité d'aIF2 $\alpha\gamma$  de *S. solfataricus* pour le Met-ARNt<sup>Met</sup> d'*E. coli* (2 nM) en présence de 1 mM GDPNP-Mg<sup>2+</sup> ou de 1 mM GDP-Mg<sup>2+</sup>.

Un facteur 80 existe donc entre les Kds des deux formes  $\alpha\gamma$ :GDPNP et  $\alpha\gamma$ :GDP de l'hétérodimère. Ce résultat est en accord avec les considérations structurales évoquées précédemment qui montrent que la formation d'un site optimal de liaison de l'ARNt nécessite un mouvement de la région des Switch au sein d'aIF2 $\gamma$ . Ce mouvement est effectivement observé entre les deux structures aIF2 $\alpha\gamma$ :GDPNP de *S. solfataricus* et aIF2 $\gamma$ :GDP de *P. abyssi* et il crée la poche de fixation du résidu méthionine de l'ARNt. Ce résultat explique donc de manière convaincante la GTP-dépendance de la liaison de l'ARNt par aIF2.

Récemment, un résultat similaire a été obtenu pour le facteur d'élongation SelB qui est responsable de l'incorporation de sélénocystéine (Sec) au sein des chaînes polypeptidiques (Leibundgut *et al.*, 2005). Ce facteur reconnaît en effet spécifiquement l'ARNt<sup>Sec</sup> qui n'est pas pris en charge par le facteur d'élongation EF1A (Rudinger *et al.*, 1996). Cet ARNt<sup>Sec</sup> est chargé avec une sélénocystéine et SelB l'incorpore au niveau d'un codon stop UAG présent au site ribosomal A si une séquence SECIS caractéristique (structure en épingle à cheveu présentant des éléments consensus) est présente en aval sur l'ARNm pour indiquer l'incorporation de la Sec. Ce facteur est homologue au facteur d'élongation EF1A, mais également à la sous-unité  $\gamma$  d'e/aIF2. SelB remplit une fonction similaire à e/aIF2 puisqu'il reconnaît un ARNt spécifique, contrairement à EF1A qui reconnaît tous les autres ARNts. Par

contre, il intervient dans la phase d'élongation. Dans le cas de SelB comme dans le cas d'aIF2 $\gamma$ , les trois structures obtenues sous forme apo, GDP et GTP présentent toutes une configuration générale qui correspond à la conformation fermée de la forme active EF1A:GTP:ARNt. Aucun changement conformationnel majeur n'est observé entre les différentes formes de SelB. Par contre, un réarrangement subtil de la région Switch II a lieu lors de la fixation du GTP et les auteurs proposent ainsi une explication pour la liaison GTP-dépendante de l'ARNt<sup>Sec</sup>.

Les changements conformationnels permettant cette modulation d'affinité entre les états GDP et GTP semblent donc ne mettre en jeu que de subtils ajustements de la position respective des différents domaines d'aIF2 $\gamma$ . Le large mouvement du domaine 1 observé dans le cas d'EF1A n'est pas retrouvé chez son homologue du démarrage de la traduction. Pour le facteur d'élongation, une telle réorganisation est supposée survenir dans les trois domaines du monde vivant (Andersen *et al.*, 2003 ; Vitagliano *et al.*, 2001). Son absence chez e/aIF2 pourrait impliquer qu'un tel mouvement ne soit pas nécessaire dans le cas du facteur de démarrage. De manière intéressante, la conformation ouverte du facteur EF1A:GDP est aussi systématiquement observée dans les complexes EF1A:EF1B (Jeppesen *et al.*, 2005 ; Andersen *et al.*, 2000 ; Wang *et al.*, 1997 ; Kawashima *et al.*, 1996). Il est donc possible que cette conformation ouverte de la forme GDP d'EF1A soit en fait liée à la présence du facteur d'échange. Dans ce cas, il serait logique de ne jamais observer cette conformation pour la protéine aIF2 $\gamma$  puisqu'il n'existe pas de facteur d'échange chez les Archées, la transition GDP  $\rightarrow$  GTP étant supposée survenir de manière spontanée.

Cette différence structurale entre e/aIF2 $\gamma$  et EF1A pourrait enfin refléter une divergence au sein des mécanismes de relargage de ces deux facteurs du ribosome. En effet, dans le cas d'eIF2, l'éjection du facteur est conditionnée par l'hydrolyse de la molécule de GTP que porte  $\gamma$  et cette hydrolyse est déclenchée par la GAP eIF5. Aucun homologue d'eIF5 n'existant chez les Archées, le ribosome pourrait lui-même activer la fonction GTPase du facteur aIF2 ou bien un autre facteur encore inconnu pourrait jouer ce rôle (voir **Chapitre II**). Dans tous les cas, un mécanisme spécifique des Archées pourrait entrer en jeu – un mécanisme nécessitant une autre réorganisation structurale d'aIF2 que celle rencontrée chez EF1A.

L'ensemble des résultats décrits dans ce **Chapitre V** a fait l'objet d'une publication dans *Structure*.



*CONCLUSION*



Le facteur hétérotrimérique e/aIF2 joue un rôle central dans le démarrage de la traduction chez les Eucaryotes et chez les Archées. Il conduit l'ARNt initiateur méthionylé jusqu'au ribosome et assure la spécificité de sélection du codon de démarrage sur l'ARNm. La structure cristallographique d'aIF2 $\gamma$  de l'archée *Pyrococcus abyssi*, précédemment résolue au laboratoire, a révélé une très forte homologie entre aIF2 $\gamma$ , qui constitue le cœur de l'hétérotrimère, et le facteur d'élongation bactérien EF1A. La sous-unité  $\gamma$  du facteur de démarrage e/aIF2 pourrait ainsi reconnaître la molécule d'ARNt selon un mode similaire à celui observé pour EF1A. Cependant,  $\gamma$  possède des caractéristiques structurales propres qui pourraient être responsables de sa spécificité d'action dans le démarrage de la traduction.

Le travail proposé avait pour objectif de mieux comprendre le mode de reconnaissance du Met-ARNt<sub>i</sub><sup>Met</sup> par aIF2 et, en particulier, de déterminer le rôle joué par chacune des sous-unités de l'hétérotrimère dans cette liaison. Pour cela, nous avons tout d'abord choisi d'étudier le facteur aIF2 de l'archée hyperthermophile *P. abyssi*. Dans un premier temps, nous avons mis en place un test *in vitro* de protection de l'ARNt contre la déacylation spontanée pour suivre l'association du facteur aIF2 au Met-ARNt<sub>i</sub><sup>Met</sup>. Ce test nous a permis d'évaluer l'importance des déterminants nucléotidiques de l'ARNt initiateur méthionylé pour sa prise en charge spécifique par aIF2. Ainsi, le résidu méthionine joue un rôle primordial dans cette reconnaissance. Les autres déterminants de l'ARNt, en particulier la paire de bases 1-72, ont également un rôle, bien que mineur, dans la spécificité de cette interaction. Nous nous sommes ensuite intéressés à la contribution de chacune des sous-unités du facteur aIF2 à la reconnaissance du Met-ARNt<sub>i</sub><sup>Met</sup>. Nous avons établi que la sous-unité  $\gamma$  seule était capable de lier spécifiquement le Met-ARNt<sub>i</sub><sup>Met</sup>. Par contre, son affinité pour l'ARNt est très réduite comparativement à celle de l'hétérotrimère. Ce résultat nous a conduit à envisager l'implication d'une autre sous-unité dans la liaison du polynucléotide. Effectivement, en présence de la sous-unité  $\alpha$ ,  $\gamma$  retrouve une affinité optimale pour l'ARNt. Par contre, l'omission de la sous-unité  $\beta$  ne semble pas modifier l'affinité du facteur pour le Met-ARNt<sub>i</sub><sup>Met</sup>. Nous avons alors entrepris une cartographie des zones d'interaction entre aIF2 $\gamma$  et aIF2 $\alpha$  par analyse mutationnelle des deux sous-unités. Cette analyse a révélé que le domaine 3 d'aIF2 $\alpha$  interagit avec le domaine 2 d'aIF2 $\gamma$  par l'intermédiaire d'une boucle idiosyncratique du domaine 2 de  $\gamma$ . L'analyse mutationnelle d'aIF2 $\gamma$  a également permis de montrer que le mode de reconnaissance de l'ARNt par aIF2 $\gamma$  était similaire à celui du facteur d'élongation bactérien EF1A. Enfin, le dimère aIF2 $\alpha_{D2-3}\gamma$  s'est révélé nécessaire et suffisant pour retrouver une affinité pour l'ARNt comparable à celle du facteur natif.

Compte tenu du rôle primordial joué par aIF2 $\alpha$  dans la reconnaissance du Met-ARNt<sub>i</sub><sup>Met</sup>, la seconde étape de ce travail a consisté à résoudre la structure complète de cette sous-unité, et en particulier celle du domaine 3 qui était alors inconnue. Les structures d'aIF2 $\alpha$  entière et de la forme  $\alpha_{D2-3}$  ont été résolues et affinées à 3.37 et 2.26 Å,

respectivement ( $R = 24.9\%$ ,  $R_{\text{libre}} = 29.0\%$  pour  $\alpha$  ;  $R = 23.6\%$ ,  $R_{\text{libre}} = 28.5\%$  pour  $\alpha_{D2-3}$ ). Ces structures ont montré que l'organisation en 3 domaines de la protéine  $\alpha$  est conservée chez les Eucaryotes et chez les Archées. Les domaines 1 et 2 de la protéine adoptent une orientation relative fixe tandis que le domaine 3 est mobile vis-à-vis de ce bloc rigide  $\alpha_{D1-2}$ . Une comparaison des séquences d'e/aIF2 $\alpha$  a montré que les zones de conservation d' $\alpha$  se situaient principalement dans le domaine 1 et dans le domaine 3 de la protéine. De plus, la recherche d'homologues structuraux pour ces deux domaines a permis de les rapprocher de protéines liant les acides nucléiques. Nous avons confirmé par des expériences de retard de migration que ces deux domaines présentent des propriétés générales de liaison des ARNs. Au vu de ces résultats, une interaction directe entre le domaine 3 d'aIF2 $\alpha$ , qui connecte la sous-unité  $\gamma$  près de la poche présumée de fixation du résidu méthionine de l'ARNt, et le polynucléotide pouvait être envisagée. D'autre part, le domaine 1 d'aIF2 $\alpha$ , qui n'interagit pas avec l'ARNt, pourrait interagir avec un autre partenaire de type ARN du démarrage de la traduction, tel que l'ARNm ou un ARN ribosomal. Enfin, la mobilité du domaine 3 au sein d'aIF2 $\alpha$  pourrait permettre le relais d'informations entre  $\gamma$  et d'autres partenaires au cours du démarrage de la traduction.

A partir de ces résultats, nous avons reporté notre intérêt sur l'hétérodimère  $\alpha\gamma$ . Les essais cristallographiques menés sur les protéines du facteur de *P. abyssi* ont conduit à l'obtention de cristaux du dimère  $\alpha_{D3}\gamma$  diffractant à 4.7 Å et une solution de remplacement moléculaire a été obtenue. Cependant, la structure du complexe n'a pu être résolue. Nous avons donc choisi d'étudier les protéines issues d'un nouvel organisme : *Sulfolobus solfataricus*. Les sous-unités séparées du facteur Ss-aIF2 ont été produites et nous avons confirmé les résultats obtenus avec les protéines de *P. abyssi* concernant le rôle des différentes sous-unités dans la liaison de l'ARNt : comme chez *P. abyssi*,  $\alpha$  est essentiel à la liaison optimale de l'ARNt initiateur méthionylé et au sein de Ss-aIF2 $\alpha$ , le domaine 3 à lui seul peut accomplir cette fonction. Toutefois, chez *S. solfataricus*, la sous-unité  $\beta$  pourrait avoir un rôle, bien que minoritaire, dans cette interaction. De plus, la sous-unité  $\gamma$  seule ne semble pas capable de lier le Met-ARNt<sup>Met</sup>. Ces différences pourraient être liées aux conditions de croissance qui diffèrent entre ces deux organismes. En effet, la température de 51°C utilisée lors de nos expériences est plus proche de la température optimale de croissance de *S. solfataricus* que de celle de *P. abyssi*.

Finalement, la résolution de la structure de l'hétérodimère aIF2 $\alpha\gamma$  de *S. solfataricus* à 3.0 Å de résolution ( $R = 22.4\%$ ,  $R_{\text{libre}} = 29.1\%$ ) a confirmé l'identité des régions des deux protéines impliquées dans l'interaction  $\alpha$ - $\gamma$  et a permis d'apporter une lumière nouvelle sur le mécanisme de reconnaissance de l'ARNt par aIF2. Au sein de cette structure, les régions Switch du domaine G d'aIF2 $\gamma$  adoptent une conformation similaire à celle observée au sein de

la forme active EF1A:GTP:Phe-ARNt<sup>Phe</sup> du facteur d'élongation, permettant d'expliquer la GTP-dépendance de la liaison du Met-ARNt<sub>i</sub><sup>Met</sup> par aIF2.

Au regard du travail accompli, de nouvelles questions sont apparues. Un nouveau modèle de positionnement de l'ARNt au sein de ce complexe  $\alpha\gamma$ , plus précis, a été réalisé. Il suggère que la sous-unité  $\alpha$  n'entrerait pas en contact direct avec le polynucléotide ARNt. A partir de ces seuls résultats, une interaction directe entre  $\alpha$  et l'ARNt ne peut cependant pas être exclue. Toutefois,  $\alpha$  pourrait contribuer au renforcement de l'affinité de  $\gamma$  pour l'ARNt par un mécanisme indirect. Notamment,  $\alpha$  pourrait stabiliser les mouvements conformationnels observés au sein de la sous-unité  $\gamma$  lors de la fixation du GTP et qui permettent de placer aIF2 $\gamma$  dans une conformation active pour lier l'ARNt. La résolution de la structure du complexe  $\alpha\gamma$  en présence d'ARNt constitue l'étape suivante dans la compréhension du rôle joué par  $\alpha$  et  $\gamma$  dans la reconnaissance du polynucléotide. L'obtention de structures de  $\gamma$  et d' $\alpha\gamma$  cocrystallisés avec différents nucléotides (GDP, GTP) pourrait également permettre de préciser la part imputable à  $\alpha$  et celle imputable au seul GTP dans la réorganisation conformationnelle de  $\gamma$  qui permet à cette sous-unité d'être en conformation active pour lier l'ARNt.

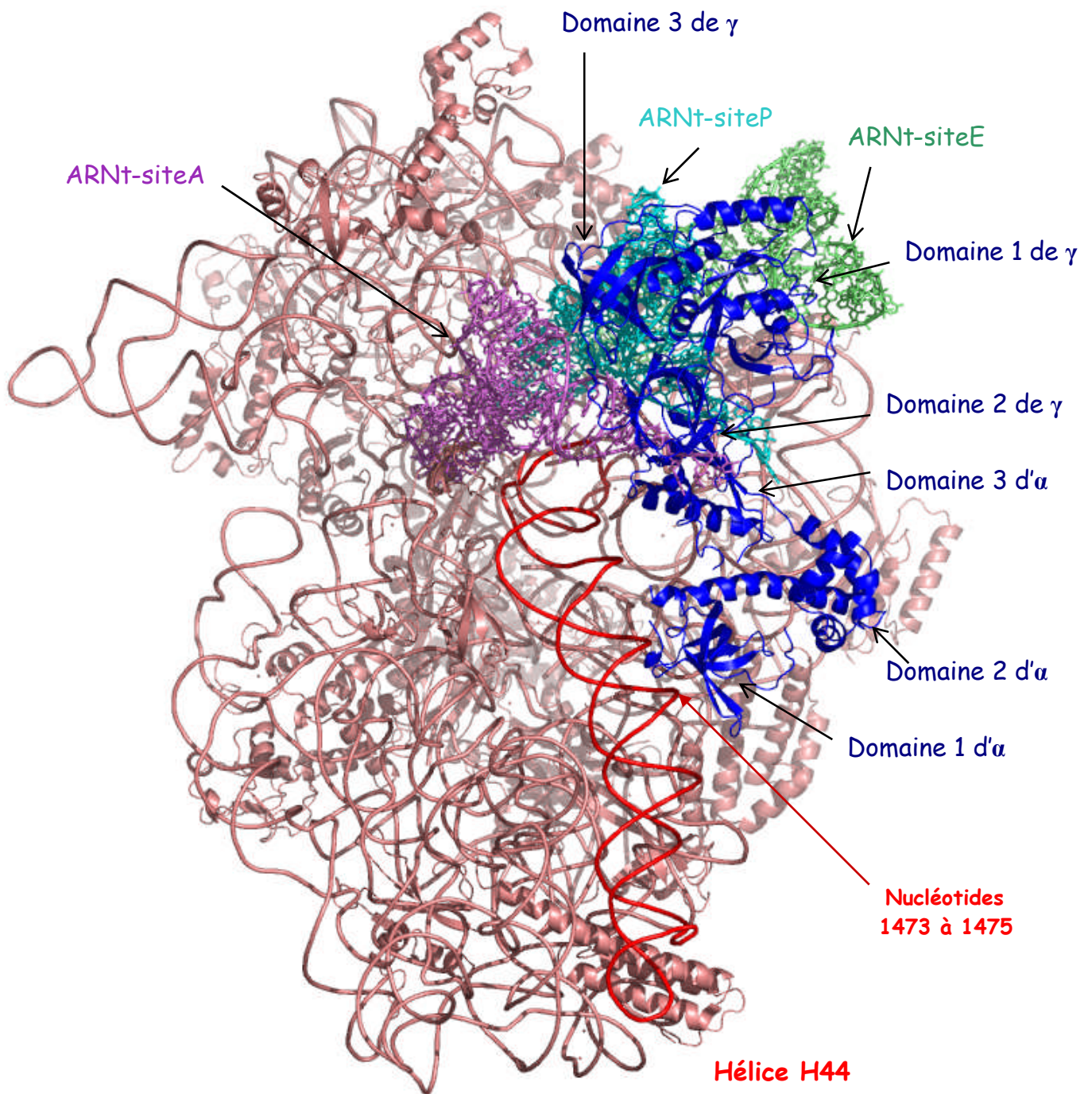
Une autre question majeure demeure : le rôle joué par la sous-unité  $\beta$ , ainsi que sa position au sein de l'hétérotrimère. Au vu des résultats que nous avons obtenus,  $\beta$  ne semble en effet pas impliquée dans la liaison du Met-ARNt<sub>i</sub><sup>Met</sup>. Chez les Eucaryotes,  $\beta$  pontait l'interaction d'eIF2 avec la GAP eIF5 et la GEF eIF2B. Chez les Archées, ces facteurs ne possèdent pas d'homologues. Dans ces conditions, le rôle joué par aIF2 $\beta$  soulève bien des interrogations. De plus, malgré l'analyse mutationnelle que nous avons entreprise sur  $\gamma$ , nous n'avons pas pu identifier la zone d'interaction entre aIF2 $\beta$  et aIF2 $\gamma$ . Le positionnement de  $\beta$  par rapport à la sous-unité  $\gamma$  et à l'ARNt reste donc à déterminer. En particulier, il a été proposé que le domaine C-terminal de  $\beta$  lie l'ARNm. Or la zone d'interaction entre  $\gamma$  et  $\beta$  est située dans le domaine central d'eIF2 $\beta$  (domaine N-terminal d'aIF2 $\beta$ ). Compte tenu de la taille relativement faible de la protéine  $\beta$ , il est difficile d'imaginer que la protéine puisse à la fois lier  $\gamma$  et l'ARNm qui se situe près de la boucle anticodon de l'ARNt, à une distance relativement grande de la sous-unité  $\gamma$  par rapport à la taille de  $\beta$ . Dans ces conditions,  $\beta$  pourrait plutôt interagir avec l'ARN ribosomal.

Finalement, toutes ces considérations posent la question du positionnement du complexe ternaire aIF2:GTP:Met-ARNt<sub>i</sub><sup>Met</sup> sur le ribosome. A partir de la structure de l'hétérodimère Ss-aIF2 $\alpha\gamma$ , il était tentant de réaliser un modèle de positionnement d' $\alpha\gamma$  au niveau du site ribosomal P. Pour cela, nous avons utilisé la structure de la sous-unité 30S de

*T. thermophilus* au sein du complexe IF1:30S (Carter *et al.*, 2001). La position des ARNts aux sites A, P et E a été obtenue à partir de la structure du ribosome 70S de *T. thermophilus* cristallisé en présence des 3 ARNts (Yusupov *et al.*, 2001). Le positionnement du complexe aIF2 $\alpha\gamma$  a été réalisé en superposant au préalable l'ARNt du site P et la molécule d'ARNt du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995), puis en superposant aIF2 $\gamma$  et EF1A (**Figure 5.27**).

Dans ce modèle, la face concave du domaine 1 d'aIF2 $\alpha$  se retrouve à proximité des nucléotides 1473 à 1475 de l'hélice H44 de l'ARN ribosomal 16S. Cette hélice H44 joue un rôle primordial dans la fonction du ribosome. En effet, les nucléotides 1492 et 1493 de cette hélice font partie du centre de décodage du site A (Clemons *et al.*, 2001). De plus, H44 se prolonge jusqu'au site P de la 30S. Ce modèle rend ainsi compte des propriétés de liaison des ARNs observées pour le domaine 1 d'aIF2 $\alpha$ . La face concave de l'OB-fold serait en contact avec l'ARNr 16S, ce qui est cohérent avec le mode de reconnaissance des ARNs par les protéines présentant un tel motif. D'autre part, ce modèle relie la sous-unité  $\alpha$  et le site ribosomal P. Il pourrait ainsi fournir une explication au rôle joué par  $\alpha$  dans la sélection du bon codon de démarrage. On peut également noter que lorsque le positionnement est réalisé sur l'ARNt présent au site ribosomal A, l'interaction du domaine 1 d'aIF2 $\alpha$  avec l'hélice H44 est toujours observée.

Finalement, ce modèle est cohérent avec l'hypothèse d'un relais de l'information entre aIF2 $\gamma$  et le ribosome lors du démarrage de la traduction. En particulier, lorsque l'appariement codon-anticodon est réalisé au site P, l'information pourrait être transmise *via* l'hélice H44 puis aIF2 $\alpha$  (par l'intermédiaire d'un mouvement d' $\alpha_{D3}$  par rapport aux domaines 1 et 2 d' $\alpha$ ) jusqu'à la sous-unité  $\gamma$  qui déclencherait alors l'hydrolyse du GTP. Ce modèle soulève de nombreuses questions particulièrement intéressantes et des études complémentaires sont nécessaires pour comprendre les mécanismes qui permettent de coupler la reconnaissance du bon codon de démarrage et l'hydrolyse du GTP porté par e/aIF2 au sein de la machinerie traductionnelle.



**Figure 5.27 :** Modèle de positionnement de l'hétérodimère aIF2 $\alpha$  de *S. solfataricus* sur la sous-unité ribosomale 30S de *T. thermophilus*. La structure de la sous-unité 30S provient de la structure du complexe IF1:30S de *T. thermophilus* (Carter *et al.*, 2001). La position des ARNts aux sites A, P et E (respectivement en violet, cyan et vert) a été déduite de la structure du complexe 70S de *T. thermophilus* (Yusupov *et al.*, 2001). Un positionnement du complexe EF1A:GDPNP:Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995) a été réalisé au préalable au niveau du site ribosomal P, par superposition des deux molécules d'ARNt. Puis, la sous-unité  $\gamma$  du complexe aIF2 $\alpha$  de *S. solfataricus* a été superposée au facteur EF1A. Le complexe Ss-aIF2 $\alpha$  est représenté en bleu foncé sur ce modèle. Enfin, l'hélice H44 de la sous-unité 30S est représentée en rouge sur la structure du ribosome (protéines et ARNr en rose pâle).



# *Matériels et Méthodes*



## **Réactifs et techniques générales de biologie moléculaire.**

Les expériences de mutagenèse dirigée et de clonage ont été réalisées par PCR sur des appareils *GeneAmp 2400* (Perkin Elmer, Wellesley, Maryland). Les oligonucléotides ont été commandés chez MWG Biotech AG (Ebersberg, Allemagne) et les désoxyribonucléotides chez Amersham (Uppsala, Suède).

Pour la production des domaines séparés des sous-unités  $\alpha$  et  $\beta$  du facteur aIF2, un protocole de PCR classique a été utilisé à partir de 5 ou 10 ng d'ADN cible, 50 pmoles de chacune des amorces oligonucléotidiques et les quatre désoxyribonucléotides (dATP, dCTP, dGTP, dTTP) à une concentration finale de 200  $\mu$ M chacun. 2 unités d'ADN polymérase thermosensible (*Taq DNA Polymerase* de Roche) ainsi que 0,4 unités de polymérase avec activité de relecture (*Pfu* de Stratagene) dans le tampon approprié ont été ajoutées dans le mélange. Les PCR ont été réalisées sur 30 cycles comprenant une phase de dénaturation à 94°C pendant 15s, une phase d'hybridation à 51°C pendant 30s et une phase d'élongation à 72°C pendant 1 min. Les produits de PCR ont ensuite été purifiés sur colonne à ADN (colonne TSK Gel DNA, TOSOHAAS, Japon).

Les mutants ponctuels et les mutants de délétion des sous-unités aIF2 $\gamma$  et aIF2 $\alpha$  ont été obtenus avec le kit *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene. Ce protocole utilise 50 ng d'ADN cible, 15 pmoles de chacune des amorces oligonucléotidiques, les quatre désoxyribonucléotides à une concentration finale de 200  $\mu$ M chacun et 2,5 unités de polymérase *Pfu Turbo* (Stratagene). Les PCR ont été réalisées sur 12, 16 ou 18 cycles selon qu'il s'agit d'une mutation ponctuelle, de la mutation d'un seul codon ou d'une insertion/délétion. Chaque cycle comporte une phase de dénaturation à 95°C pendant 30s, une phase d'hybridation à 50°C pendant 30s et une phase d'élongation à 68°C pendant 10min. Les produits de PCR ont alors été digérés par 10 unités de *DpnI* (New England Biolabs, Beverly, Massachusetts) durant 1 heure à 37°C. Cette enzyme de restriction coupe spécifiquement l'ADN méthylé parental non muté. Les plasmides mutés ont alors été directement transformés dans les souches d'*Escherichia coli* XL1 Blue.

Les enzymes de restriction utilisées pour les digestions enzymatiques proviennent de chez Roche (Mannheim, Allemagne), New England Biolabs (Beverly, Massachusetts) ou Amersham (Uppsala, Suède). Les réactions de ligation ont été réalisées avec l'ADN ligase du phage T4 (GIBCO BRL).

Les électrophorèses d'ADN ont été réalisées sur des gels à 1% (p/v) d'agarose LE (Roche, Mannheim, Allemagne), en présence de 0,3 mg/L de bromure d'éthidium (AMRESCO).

Les transformations de bactéries électrocompétentes ont été effectuées à l'aide d'un électroporateur *Gene Pulser* (Biorad, Hercules, Californie). Les bactéries électrocompétentes ont été préparées suivant la méthode de Dower *et al.* (1988). Les différentes souches d'*E.coli* utilisées sont répertoriées dans le tableau suivant :

Souche	Génotype	Référence
JM101Tr	$\Delta(lac-pro) supE thi recA56 srl-300::Tn10$ (F' <i>traD36 proAB lacI<sup>f</sup> lacZ</i> $\Delta$ M15)	Hirel <i>et al.</i> , 1988
XL1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1</i> <i>lac</i> [F' <i>proAB lacI<sup>f</sup> lacZ</i> $\Delta$ M15 Tn10 (Tet <sup>R</sup> )]	Stratagene
BL21(DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> $\lambda$ (DE3)	Novagen
Rosetta(DE3)pLacI	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> $\lambda$ (DE3) pLacIRARE (Cam <sup>R</sup> )	Novagen
B834(DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> $\lambda$ (DE3) <i>met</i>	Novagen

Les minipréparations d'ADN plasmidique ont été réalisées suivant le protocole de Birnboim et Doly (1979). Les préparations de plasmides ont été obtenues à l'aide du système *QIAfilter* (QIAGEN). Les différents plasmides utilisés lors de ce travail sont indiqués dans le tableau suivant :

Plasmide	Génotype	Référence
pET-3alpa	Amp <sup>R</sup> <i>lacI<sup>f</sup></i>	Guillon <i>et al.</i> , 2005
pET-15blpa	Amp <sup>R</sup> <i>lacI<sup>f</sup></i>	Guillon <i>et al.</i> , 2005
pSJS1240	Spec <sup>R</sup> <i>lacI<sup>f</sup></i>	
pSBETa	Kan <sup>R</sup> <i>lacI<sup>f</sup></i>	Schenk <i>et al.</i> , 1995

Le séquençage de l'ADN a été réalisé sur un séquenceur automatique *Long ReadIR 4200* (Li-Cor, Lincoln, Nebraska). Les gels de migration ont été préparés suivant les spécifications du constructeur. Le protocole de préparation des réactions de séquence est dérivé de la méthode de terminaison de chaîne (Sanger *et al.*, 1982). Les échantillons ont été

préparés par PCR à l'aide du kit de séquençage fourni par Epicentre Technologies (TEBU, Epicentre Technologies, Madison, Wisconsin). Les amorces, marquées à leur extrémité 5' par des fluorophores excitables dans l'Infra-Rouge (*IRDye 700* et *IRDye 800*), sont commandées chez MWG Biotech AG (Ebersberg, Allemagne).

Les alignements de séquences ont été réalisés à l'aide des programmes *Clustal X* (Thompson *et al.*, 1997) et *Clustal W* ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw) ; Chenna *et al.*, 2003). Les manipulations de séquences nucléiques ou protéiques ont été effectuées avec *DNAid* (Dardel et Bensoussan, 1988) ou *DNAstrider* (Marck, 1988).

Les gels d'électrophorèse de protéines en conditions dénaturantes (SDS-PAGE) ont été réalisés à l'aide du système *Protean II* de Biorad (Hercules, Californie). La préparation des gels et des tampons de migration suit les recommandations du fabricant. Le *Coomassie Brilliant Blue* provient de chez Research Organics (Cleveland, Ohio). Les colorations au nitrate d'argent ont été réalisées grâce au kit *Silver Stain Plus* de Biorad (Hercules, Californie).

Les gels d'électrophorèse pour protéines basiques en conditions non dénaturantes ont été réalisés avec le système *Protean II* de Biorad (Hercules, Californie). Les gels de migrations ont été préparés selon le même protocole que pour la SDS-PAGE mais en remplaçant le tampon de séparation par un tampon KOH/acide acétique de pH 4.3, le tampon de dépôt par un tampon KOH/acide acétique de pH 6.8 et en omettant le SDS dans le mélange. Le tampon de migration est un tampon acide acétique/ $\beta$ -alanine de pH 4.4.

Les mesures spectrophotométriques UV-visible ont été effectuées sur un appareil *UV2101PC* (Shimadzu).

Les colonnes de chromatographie (*Q-Hiload*, *S-Hiload*, *Superdex 75*, *Superdex 200*, *Phenyl-Superose*, *Mono Q*, *Mono S*) utilisées pour la purification par FPLC des différentes protéines étudiées proviennent de chez Amersham Pharmacia Biotech (Uppsala, Suède). Les différentes purifications ont été réalisées soit grâce au *Liquid Chromatography Controller LCC-500* (Amersham Pharmacia Biotech, Uppsala, Suède) qui peut être géré informatiquement par le programme *FPLC Manager* (Amersham Pharmacia Biotech, Uppsala, Suède), soit à l'aide du système chromatographique *BioLogic DuoFlow* de Biorad (Hercules, Californie). Les protéines purifiées ont été concentrées par centrifugation sur membrane en utilisant des unités de centrifugation *Vivaspin 20 mL*, *10 000 MWCO* (Vivascience AG, Hannovre, Allemagne) ou des unités de centrifugation *Amicon Ultra 4 mL*, *10 000* ou *30 000 MWCO* (Millipore).

La [<sup>35</sup>S]-Méthionine, la [<sup>14</sup>C]-Méthionine et la [<sup>3</sup>H]-Valine utilisées pour aminoacyler les ARNts proviennent de chez Nen Life Science (Perkin Elmer, Wellesley, Maryland). Les mesures de radioactivité ont été effectuées sur un compteur à scintillation *LS 6500* (Beckman Coulter).

Les matrices d'échantillonnage *Crystal Screen 1*, *Crystal Screen 2*, *Natrix*, *SaltRX*, *Index*, *Grid Screen Ammonium Sulfate*, *Grid Screen PEG 6000*, *Grid Screen MPD* et *Grid Screen NaCl* (Hampton Research, Laguna Niguel, Californie) ont été employées pour la recherche des conditions de cristallisation des différentes protéines. Les différents essais de cristallisation ont été réalisés via la technique de la goutte suspendue à l'aide d'un robot *Genesis RSP 100* (TECAN, Männedorf, Suisse). Les produits utilisés pour l'amélioration des conditions de cristallisation proviennent de chez Hampton Research (Laguna Niguel, Californie), Fluka ou SIGMA-Aldrich.

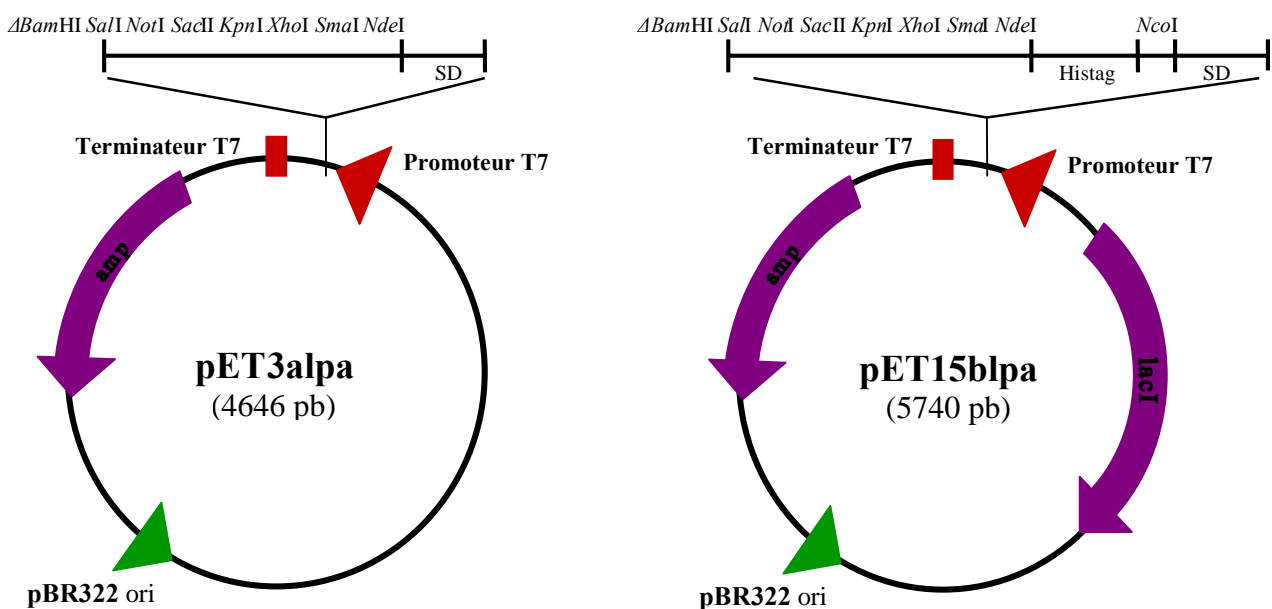
Les figures cristallographiques ont été tracées avec les logiciels *SETOR* (Evans, 1993) et *PyMOL* ([www.pymol.org](http://www.pymol.org), DeLano Scientific LLC, San Carlos, Californie).

Enfin, tous les produits chimiques courants ont été commandés chez Merck, Fluka ou Sigma-Aldrich.

## **I. CARTOGRAPHIE FONCTIONNELLE DU FACTEUR aIF2.**

### **A. Production et purification des différentes sous-unités du facteur hétérotrimérique aIF2 de *P. abyssi*.**

Les ADNs codant pour chacune des 3 sous-unités du facteur hétérotrimérique aIF2 ont été clonés dans le vecteur pET3alpa entre les sites de restriction de *NdeI* et *XhoI* du fragment « lpa » additionnel, donnant ainsi naissance aux vecteurs pET3aαlpa, pET3aβlpa et pET3aγlpa. Ce clonage a également été reproduit avec le vecteur pET15blpa (plasmides construits pET15bαlpa, pET15bβlpa, pET15bγlpa) pour pouvoir éventuellement exprimer les protéines correspondantes avec un Tag poly-Histidine à leur extrémité N-terminale. Les deux vecteurs de clonage sont représentés ci-dessous avant insertion des fragments d'ADN :



**Figure 6.1 :** Les plasmides pET3alpa et pET15blpa.

Les vecteurs pET3aαlpa, pET3aβlpa et pET3aγlpa portant les gènes natifs codant pour les 3 sous-unités d'aIF2 ont été utilisés pour transformer les cellules d'*E. coli* BL21(DE3). Pour l'expression d'aIF2<sub>γ</sub>, les cellules BL21 ont été co-transformées avec le pET3aγlpa et avec le plasmide pSJS1240 qui surexprime deux ARNts permettant le décodage des codons rares AGA, AGG (codant pour Arg) et ATA (codant pour Ile). Ces codons rares sont rencontrés plus fréquemment sur les ARNms d'Archées. Pour l'expression d'aIF2<sub>α</sub> et d'aIF2<sub>β</sub>, les cellules BL21 ont été co-transformées avec le plasmide pSBETa qui surexprime le tRNA<sub>4</sub><sup>Arg</sup> permettant de lire les codons AGA et AGG. Dans tous les cas, ces cellules sont mises en culture à 37°C dans 1 L de milieu 2xTY (Miller, 1992) contenant 100

$\mu\text{g/mL}$  d'ampicilline et  $30 \mu\text{g/mL}$  de spectinomycine pour les cellules co-transformées avec le pSJS1240 (respectivement  $50 \mu\text{g/mL}$  d'ampicilline et  $100 \mu\text{g/mL}$  de kanamycine pour les cellules co-transformées avec le pSBETa). La surexpression de la protéine d'intérêt est alors induite par ajout d'IPTG à une concentration finale de  $1 \text{ mM}$  lorsque la densité optique de la culture à  $650 \text{ nm}$  ( $\text{DO}_{650 \text{ nm}}$ ) atteint  $1,6$ . L'induction est poursuivie durant  $6$  heures à température ambiante. Puis les cultures sont centrifugées à  $8000 \text{ g}$  pendant  $30$  minutes et les culots sont congelés et conservés à  $-20^\circ\text{C}$  jusqu'à utilisation.

### **1. Purification de Pa-aIF2 $\gamma$ .**

Le culot de cellules correspondant à  $1 \text{ L}$  de culture est resuspendu dans environ  $80 \text{ mL}$  de tampon  $A_1$  ( $10 \text{ mM}$  MOPS pH  $6,7$ ,  $10 \text{ mM}$  2-mercaptoethanol,  $500 \text{ mM}$  NaCl). Les cellules bactériennes sont alors lysées par sonication aux ultrasons pendant  $6 \text{ min}$ . ( $50\%$  du temps actif) à  $4^\circ\text{C}$ . Après centrifugation ( $15300 \text{ rpm}$ ,  $10$  minutes,  $4^\circ\text{C}$ ), le surnageant est chauffé à  $80^\circ\text{C}$  pendant  $10$  minutes pour précipiter les protéines bactériennes thermosensibles. Ces dernières sont alors séparées des protéines thermostables par centrifugation ( $15300 \text{ rpm}$ ,  $10$  minutes,  $4^\circ\text{C}$ ). Le surnageant est alors précipité au sulfate de streptomycine (à une concentration finale de  $1\%$  w/v) pour faire précipiter les acides nucléiques. Enfin, après centrifugation ( $15300 \text{ rpm}$ ,  $10$  minutes,  $4^\circ\text{C}$ ), le lysat soluble restant est précipité au sulfate d'ammonium à  $80\%$  de saturation. Le culot obtenu après centrifugation est resuspendu dans  $5 \text{ mL}$  de tampon  $A_1$  et dialysé deux fois  $30$  minutes contre  $500 \text{ mL}$  de tampon  $A_1$  auquel a été ajouté du phenylméthylsulfonyl fluorure (PMSF) à une concentration finale de  $0,1 \text{ mM}$ . Cette étape de dialyse permet d'éliminer le sulfate d'ammonium restant. La protéine aIF2 $\gamma$  est ensuite purifiée en  $4$  étapes.

Le dialysat obtenu est chargé sur une colonne échangeuse d'anions Q-Hiload ( $1,6 \times 20\text{cm}$ ) équilibrée en tampon  $A_1$  à  $2,5 \text{ mL}\cdot\text{min}^{-1}$ . La protéine n'est pas retenue sur cette colonne qui permet d'éliminer une grande partie des acides nucléiques. Les fractions de tête contenant aIF2 $\gamma$  sont rassemblées et mises à dialyser  $1$  heure contre  $1 \text{ L}$  de tampon  $A_2$  ( $10 \text{ mM}$  MOPS pH  $6,7$ ,  $10 \text{ mM}$  2-mercaptoethanol,  $100 \text{ mM}$  NaCl).

La seconde étape consiste en une colonne échangeuse de cations S-Sépharose ( $1,1 \times 4 \text{ cm}$ ). Le dialysat précédent est chargé à  $2 \text{ mL}\cdot\text{min}^{-1}$  sur cette colonne équilibrée en tampon  $A_2$ . Là encore, aIF2 $\gamma$  passe au travers de la colonne. Les fractions contenant la protéine sont rassemblées.

L'échantillon est alors injecté à  $1 \text{ mL}\cdot\text{min}^{-1}$  sur une colonne échangeuse d'anions Mono-Q ( $0,5 \times 5 \text{ cm}$ ) équilibrée en tampon  $A_2$ . La protéine n'est pas retenue sur cette

colonne. Les fractions contenant aIF2 $\gamma$  sont regroupées et l'échantillon est concentré par centrifugation sur membrane *Vivaspin 20 mL, 30 000 MWCO* jusqu'à un volume inférieur à 5 mL.

Pour finir, l'échantillon est chargé à 0.4 mL.min<sup>-1</sup> sur un tamis moléculaire Superdex 75 (1,6 x 60 cm) équilibré en tampon A<sub>3</sub> (10 mM MOPS pH 6.7, 10 mM 2-mercaptoethanol, 200 mM NaCl). La protéine est éluée sous forme d'un pic homogène. Les fractions correspondantes sont rassemblées et concentrées par centrifugation sur membrane *Amicon-Ultra 4 mL, 30 000 MWCO* jusqu'à atteindre une concentration de l'ordre de 8-10 mg/mL. De l'azide de sodium à 0,02% final est ajouté pour assurer la stabilité de la protéine. Du GDPNP et du MgCl<sub>2</sub> à une concentration finale de 1 mM sont également ajoutés pour empêcher la protéolyse de la sous-unité  $\gamma$  au niveau du domaine de fixation du GTP. La protéine est alors conservée à 4°C. En moyenne, 4 mg de protéine purifiée sont obtenus à partir d'un litre de culture.

## **2. Purification de Pa-aIF2 $\beta$ .**

Le même protocole de purification que celui de la sous-unité  $\gamma$  est appliqué. Cependant, seul l'azide de sodium est ajouté à l'échantillon purifié pour la stabilité de la protéine. 20 mg de protéine sont obtenus à partir d'un litre de culture.

## **3. Purification de Pa-aIF2 $\alpha$ .**

La purification de la sous-unité  $\alpha$  suit le même protocole que celle de la sous-unité  $\gamma$  jusqu'à la première colonne chromatographique. Après la Q-Hiload, les fractions rassemblées contenant aIF2 $\alpha$  sont dialysées contre du tampon A<sub>4</sub> (10 mM MOPS pH 6.7, 10 mM 2-mercaptoethanol, 300 mM NaCl) auquel a été ajouté du PMSF à une concentration finale de 0.1 mM. Le dialysat est alors chargé à 2 mL.min<sup>-1</sup> sur une colonne échangeuse de cations S-Sépharose (1,1 x 4 cm) équilibrée en tampon A<sub>4</sub>. Après le passage des fractions de tête, la protéine est éluée par un gradient de 300 à 800 mM en NaCl en 20 minutes (25 mM/minute). Les fractions contenant aIF2 $\alpha$  sont alors rassemblées, concentrées puis injectées à 0.4 mL.min<sup>-1</sup> sur un tamis moléculaire Superdex 75 (1,6 x 60 cm) équilibré en tampon A<sub>1</sub>. La protéine, obtenue sous forme d'un pic homogène, est concentrée par centrifugation sur membrane *Amicon-Ultra 4 mL, 30 000 MWCO* jusqu'à atteindre une concentration de l'ordre de 8-10 mg/mL. 20 mg de protéine sont obtenus à partir d'un litre de culture.

#### **4. Purification des hétérodimères $\alpha\gamma$ , $\beta\gamma$ et de l'hétérotrimère $\alpha\beta\gamma$ .**

Pour chaque complexe envisagé, les sous-unités purifiées correspondantes ont été mélangées en proportions telles que la sous-unité  $\gamma$  est toujours en défaut. Le mélange est alors chargé à  $0.4 \text{ mL}\cdot\text{min}^{-1}$  sur un tamis moléculaire Superdex 75 ou Superdex 200 (1,6 x 60 cm) équilibré en tampon A<sub>3</sub>. Le complexe est élué sous forme d'un pic homogène et l'excès d' $\alpha$  et/ou de  $\beta$  élue sous forme d'un second pic nettement séparé du premier. Les fractions relatives au complexe sont alors rassemblées et concentrées à 8-10 mg/mL finaux.

#### **B. Mesure de l'affinité d'aIF2 pour l'ARN de transfert.**

L'ARNt<sub>f</sub><sup>Met</sup>, l'ARNt<sub>m</sub><sup>Met</sup> (élongateur), l'ARNt<sub>i</sub><sup>Val</sup> d'*E. coli* et leurs dérivés ont été produits selon le protocole établi (Guillon *et al.*, 1992 ; Meinnel and Blanquet, 1995). Le gène codant pour l'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi* a été construit par assemblage d'oligonucléotides et inséré dans le vecteur pBSTNAV. Ces différents ARNts ont été purifiés par sur une colonne chromatographique échangeuse d'anions (Meinnel *et al.*, 1988).

#### **1. Préparation d'ARNts aminoacylés avec de la [<sup>35</sup>S]-Met ou de la [<sup>3</sup>H]-Val.**

Les réactions de méthionylation sont réalisées dans un volume de 100  $\mu\text{L}$  dans du tampon 20 mM Tris-HCL pH 7.5, 10 mM 2-mercaptoéthanol, 0.1 mM EDTA, 150 mM KCl, 7 mM MgCl<sub>2</sub>, 2mM ATP en présence de 10  $\mu\text{M}$  d'ARNt et de 25  $\mu\text{M}$  d'une dilution isotopique de [<sup>35</sup>S]-Met (10000 dpm/pmole). Le mélange est alors incubé 10 minutes à 25°C en présence de methionyl-ARNt synthétase M547 d'*E.coli* (Mellot *et al.*, 1989) à une concentration finale de 1  $\mu\text{M}$ . Après 10 minutes d'incubation, 1  $\mu\text{L}$  de la solution est prélevé et introduit dans 3 mL du mélange d'arrêt 5% acide trichloroacétique, 0.5% méthionine. 20  $\mu\text{L}$  d'une solution d'ARN totaux de levure à 4 g.L<sup>-1</sup> sont ajoutés pour faciliter la précipitation de l'ARNt. Le tout est homogénéisé puis filtré sur des filtres en microfibrilles de verre GF/C (Whatman). La radioactivité adsorbée est mesurée dans un compteur à scintillation pour vérifier le taux d' aminoacylation de l'ARNt. Une réaction d' aminoacylation de contrôle est réalisée en présence de [<sup>14</sup>C]-Met (130 dpm/pmole). Cette réaction permet de mesurer la quantité d'ARNt effectivement aminoacylé sans être gêné par la variation de radioactivité spécifique de la [<sup>35</sup>S]-Met. Le reste du mélange est précipité à l'éthanol en présence de 0.3 M d'acétate de sodium pH 5.5. Les solutions précipitées sont ainsi conservées à -20°C.

Les réactions d' aminoacylation avec la [<sup>3</sup>H]-Val (10000 dpm/pmole) sont réalisées suivant le même protocole à la seule différence que le tampon d' aminoacylation ne contient

pas de KCl. La [<sup>3</sup>H]-Val est utilisée à une concentration de 25 µM et l'aminacylation est réalisée par la ValRS d'*E.coli* (Brevet *et al.*, 1989).

Avant utilisation, les stocks d'ARNt aminoacylé sont centrifugés (15300 rpm, 4°C, 20 min.) et le culot est repris dans de l'eau. L'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi* est repris dans un tampon 10 mM HEPES pH 8.0, 10 mM MgCl<sub>2</sub>, 200 mM KCl, 0.1 mM EDTA. La présence de Mg<sup>2+</sup> est nécessaire à son repliement correct.

## **2. Détermination de la constante de dissociation des complexes aIF2:ARNt.**

Cette détermination est basée sur la capacité d'aIF2 à protéger l'ARNt aminoacylé de la déacylation spontanée. La mesure de la constante de dissociation (Kd) des différents complexes est réalisée en suivant la cinétique de déacylation des ARNts étudiés en présence de concentrations croissantes d'aIF2.

Les réactions sont réalisées dans un volume de 150 µL dans un tampon déacylant contenant 20 mM HEPES-NaOH pH 8.0, 100 mM KCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mM EDTA, 0.2 mg/mL BSA (Bovine Serum Albumine, ROCHE), 5 mM GDPNP en présence de 80 nM de l'ARNt aminoacylé étudié. La concentration en aIF2 natif (ou assemblage dérivé) varie de 10 nM à 80 µM selon la valeur du Kd à mesurer. Pour chaque concentration en aIF2, la réaction est démarrée en ajoutant l'ARNt et le mélange est incubé à 51°C. Pour déterminer la vitesse de déacylation de l'ARNt, des aliquots de 25 µL sont prélevés à différents temps (0, 5, 10, 15, 20 et 30 minutes) et précipités dans 3 mL d'une solution d'acide trichloroacétique 5%, 0.5% méthionine en présence de 80 µg d'ARN de levure. La courbe de déacylation ainsi obtenue peut alors être modélisée par une exponentielle et le programme MC-Fit (Dardel, 1994) permet d'en déduire la constante de temps correspondante. Une courbe d'association-dissociation représentant les constantes de temps ainsi mesurées en fonction des différentes concentrations en aIF2 est alors tracée. Le programme MC-Fit permet de déduire le Kd correspondant au complexe aIF2:ARNt étudié.

## **C. Obtention des mutants de la protéine aIF2γ de *P. abyssi*.**

### **1. Obtention des mutants γ<sub>AL1</sub>, γ<sub>AL2</sub>, γ<sub>AZm</sub>, γ<sub>C60A</sub> et γ<sub>C72A</sub>.**

La première série de mutants d'aIF2γ a été obtenue par PCR classique en deux étapes successives à partir du vecteur cible pET3aγ1pa.

Dans un premier temps, pour chaque mutant, deux réactions de PCR distinctes ont été réalisées de façon à amplifier les régions en amont et en aval de la mutation. Pour chaque réaction, deux amorces ont été utilisées :

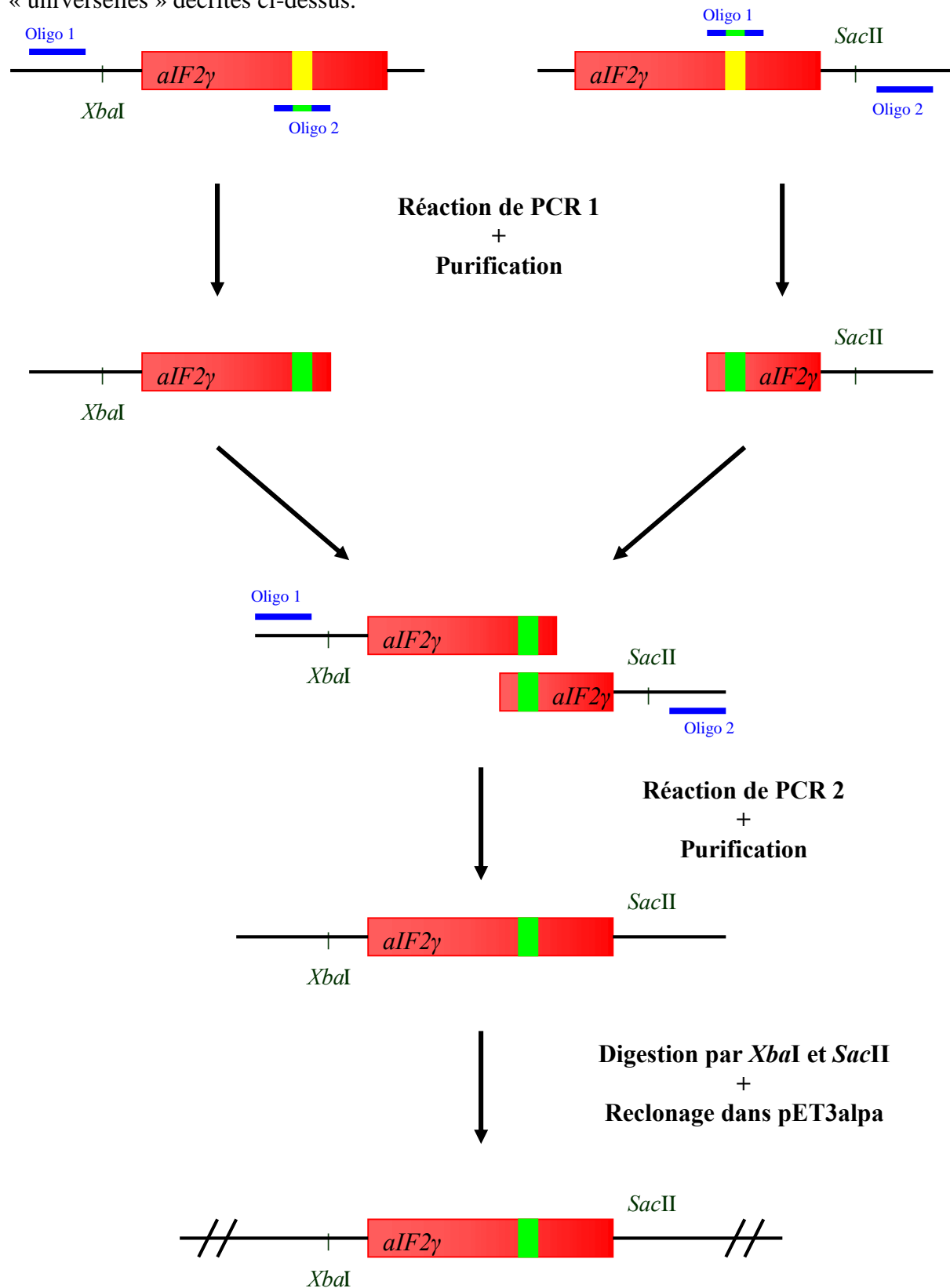
- une amorce dite « universelle » qui s'hybride à la séquence en amont ou en aval du gène d'aIF2 $\gamma$  sur le vecteur pET3alpa
- une amorce « interne », qui s'hybride sur le gène codant pour aIF2 $\gamma$ , de part et d'autre de la zone à muter et qui contient la mutation à effectuer.

Les amorces « universelles » ont été choisies de manière à ce que les fragments synthétisés contiennent en plus un site de restriction unique sur toute la séquence du pET3alpa. Ainsi, l'amorce qui permet de synthétiser le fragment amont par rapport à la zone à muter se trouve légèrement en amont du site de restriction de *Xba*I (site unique en position 4052) sur le pET3alpa. L'amorce qui permet de synthétiser la zone aval se trouve légèrement en aval du site de restriction de *Sac*II (site unique en position 5353). Pour les mutants de délétion, les résidus supprimés ont été choisis de manière à pouvoir introduire un résidu glycine (codon GGT) entre les deux résidus délimitant la zone coupée. Les fragments de PCR ainsi obtenus ont été purifiés de façon à servir de matrice pour la seconde étape. Les différentes amorces « internes » utilisées pour chaque mutant ainsi que les deux amorces « universelles » sont récapitulées dans le **Tableau 6.2**.

Mutation	Amorces	Séquence
Délétion de la boucle 221-227 $\gamma_{\Delta L1}$	Fragment amont	5' AAGCTTTGACGTCAATGGTAAGCTAGTAGGTGGCG 3'
	Fragment aval	5' CGCCACCTACTAGCTTACCATTGACGTCAAAGCTT 3'
Délétion de la boucle 256-265 $\gamma_{\Delta L2}$	Fragment amont	5' GATAAGGCCGGGCGTTGGTGAGCCGATAACGACCG 3'
	Fragment aval	5' CGGTCGTTATCGGCTCACCAACGCCCGCCTTATC 3'
Délétion du domaine du Zinc (58-81) $\gamma_{\Delta Zn}$	Fragment amont	5' TGCGGATGCCGAGATAGGTGTGAGGAGGGTTTCGT 3'
	Fragment aval	5' ACGAAACCCTCCTCACACCTATCTCGGCATCCGCA 3'
Mutant ponctuel $\gamma_{C60A}$	Fragment amont	5' AGATAAGGAGAGCTCCAAACTGCGG 3'
	Fragment aval	5' CCGCAGTTTGGAGCTCTCCTTATCT 3'
Mutant ponctuel $\gamma_{C72A}$	Fragment amont	5' CTTCCCCAGTGGCTCCCTACTGCGG 3'
	Fragment aval	5' CCGCAGTAGGGAGCCACTGGGGAAG 3'
Amorce universelle pour le fragment amont		5' GACTCACTATAGGGAGACCACAACGG 3'
Amorce universelle pour le fragment aval		5' TCCTTTCGGGCTTTGTTAGCAGCCG 3'

**Tableau 6.2 :** Amorces utilisées pour synthétiser la première série de mutants d'aIF2 $\gamma$  de *P. abyssi*. Pour les mutants ponctuels, la mutation introduite dans l'amorce est indiquée en gras. Pour les mutants de délétion, le résidu glycine introduit à la place des résidus supprimés est souligné.

Dans un second temps, pour chaque mutation, les deux fragments d'ADN mutés ont été mélangés en quantités équimolaires et amplifiés par PCR en utilisant les deux amorces « universelles » décrites ci-dessus.



**Figure 6.3 :** Principe de la mutagenèse dirigée par PCR de chevauchement.

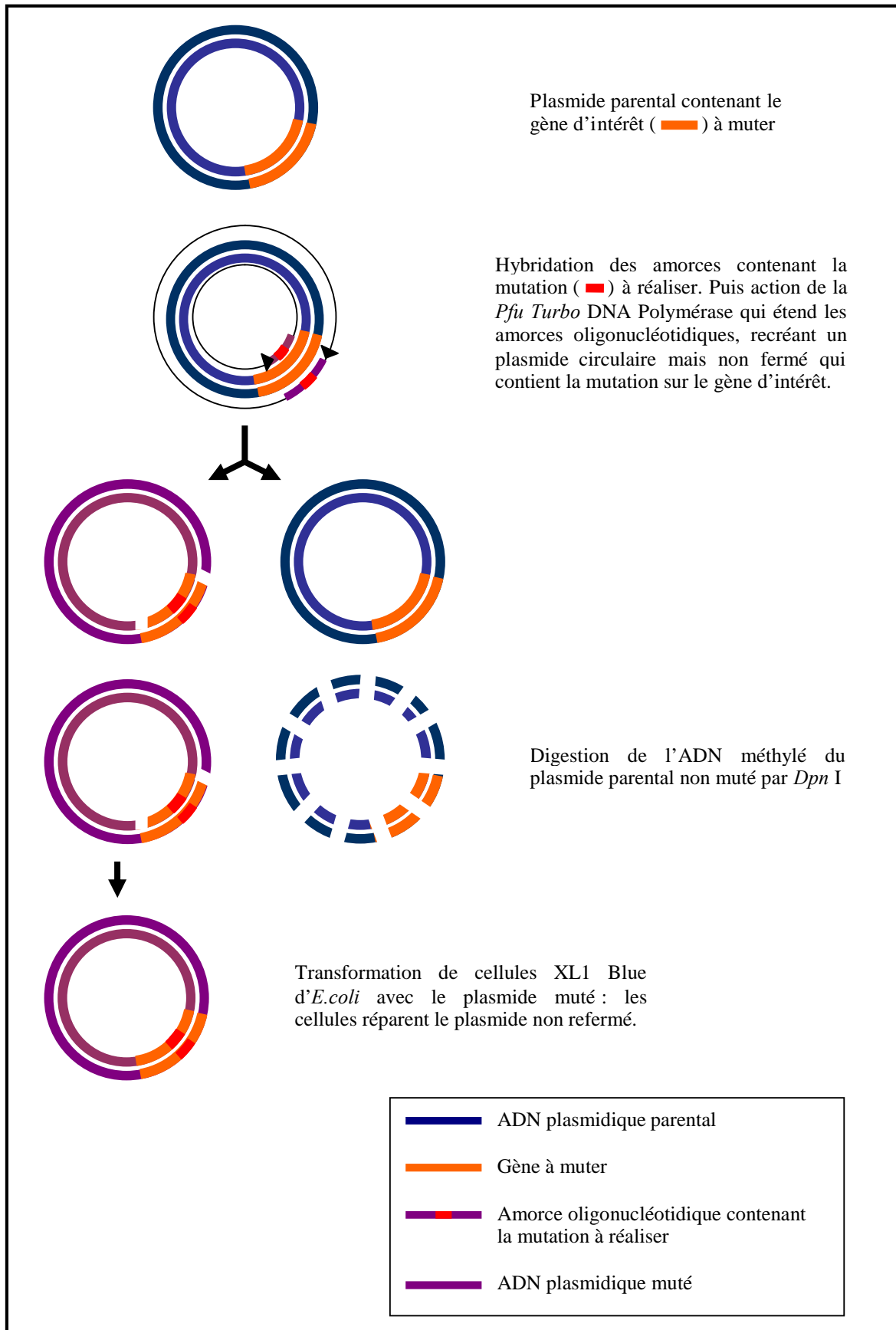
Après purification de ces nouveaux produits de PCR et digestion par les enzymes de restriction *XbaI* et *SacII*, les fragments d'ADN ont été clonés dans le plasmide pET3alpa, digéré par les mêmes enzymes, pour former les plasmides pET3aγlpa portant les différentes mutations. Ces plasmides ont été séquencés pour vérifier la présence effective des mutations souhaitées. La **Figure 6.3** détaille le principe de la méthode.

## 2. Obtention des mutants $\gamma_{\Delta L3}$ , $\gamma_{\Delta L4}$ , $\gamma_{C63H}$ , $\gamma_{P178A}$ et $\gamma_{N186A}$ .

La seconde série de mutants de la sous-unité aIF2 $\gamma$  a été obtenue à partir du vecteur cible pET3aγlpa par la méthode *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene. A partir de deux amorces oligonucléotidiques qui portent la mutation à réaliser et qui s'hybrident sur chacun des brins de l'ADN cible, de part et d'autre de la zone à muter, ce système permet de reconstruire entièrement le plasmide de départ. Il fait appel à une polymérase avec activité exonucléase 5'→3' et 3'→5', la *Pfu Turbo*, qui présente par conséquent un très haut degré de fidélité, ce qui limite les mutations spontanées non désirées pouvant survenir lors de la réaction de PCR. Le mélange réactionnel ainsi obtenu par PCR contient l'ADN cible ainsi que l'ADN muté amplifié et recircularisé mais non fermé. Ce mélange est alors digéré par l'enzyme de restriction *DpnI* qui va cliver préférentiellement l'ADN méthylé, c'est-à-dire l'ADN de départ non modifié. Ainsi, seul l'ADN muté est intact et il peut être directement utilisé pour transformer des cellules XL1 Blue d'*Escherichia coli*. Le principe détaillé de la méthode est repris dans le schéma récapitulatif de la **Figure 6.5**. Les 5 mutants construits grâce à cette méthode ainsi que les amorces oligonucléotidiques utilisées dans chaque cas sont détaillés dans le **Tableau 6.4**.

Mutation	Amorces	Séquence
Déletion de la boucle 335-341 $\gamma_{\Delta L3}$	amont	5' GTTCACTTACTTGAGAGGGG <b>ICT</b> TAAAGTCGAACCCATA 3'
	aval	5' TATGGGTTCGACTTTAAG <b>ACC</b> CCTCTCAAGTAAGTGAAC 3'
Déletion de la boucle 394-402 $\gamma_{\Delta L4}$	amont	5' GATAGGGTTGCAATAAG <b>CGG</b> TATAGGATACGGAATAATC 3'
	aval	5' GATTATTCCGTATCCTATA <b>CCG</b> CCTATTGCAACCCTATC 3'
Mutant ponctuel $\gamma_{C63H}$	amont	5' ATAAGGAGATGTCCAAACC <b>ACG</b> GTAGGTATTCAACTTCC 3'
	aval	5' GGAAGTTGAATACCTACC <b>GTG</b> GTTTGGACATCTCCTTAT 3'
Mutant ponctuel $\gamma_{P178A}$	amont	5' GAGAACGCTCCAATAATT <b>GCT</b> ATCTCAGCTTTGCATGGG 3'
	aval	5' CCCATGCAAAGCTGAGATAGCAATTATTGGAGCGTTCTC 3'
Mutant ponctuel $\gamma_{N186A}$	amont	5' TCAGCTTTGCATGGGG <b>CCG</b> CTATAGACGTCCTAGTCAAG 3'
	aval	5' CTTGACTAGGACGTCTATA <b>GCG</b> CCCCATGCAAAGCTGA 3'

**Tableau 6.4 : Amorces utilisées pour synthétiser la seconde série de mutants d'aIF2 $\gamma$  de *P. abyssii*.** Pour les mutants ponctuels, la mutation introduite dans l'amorce est indiquée en gras. Pour les mutants de déletion, le résidu glycine introduit à la place des résidus supprimés est souligné.



**Figure 6.5 :** Principe des expériences de mutagenèse dirigée par la méthode *QuikChange*<sup>TM</sup> *Site-Directed Mutagenesis* de Stratagene.

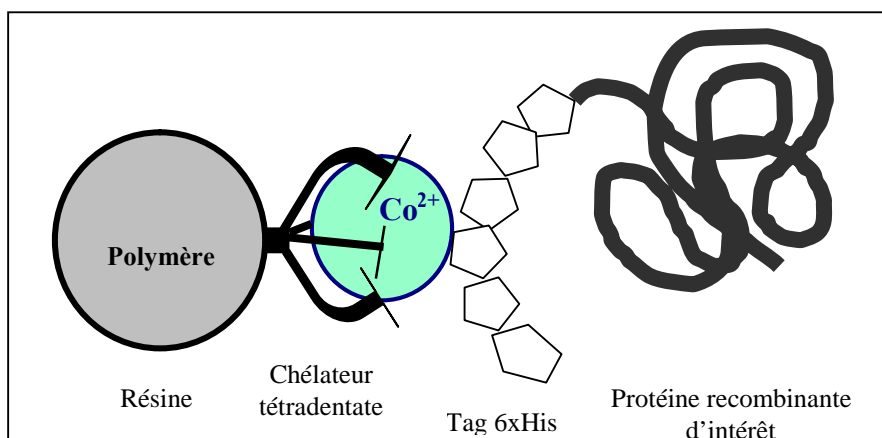
### 3. Production des mutants de Pa-aIF2 $\gamma$ .

L'expression des différents mutants d'aIF2 $\gamma$  est réalisée par co-transformation de cellules BL21 avec le plasmide pET3a $\gamma$ lpa portant l'insert muté et avec le plasmide pSJS1240. Pour l'expression de la version d'aIF2 $\alpha$  fusionnée à un hexapeptide 6xHis, c'est le vecteur pET15b $\alpha$ lpa qui a été utilisé avec le plasmide pSBETa pour co-transformer les cellules BL21. Le protocole est ensuite identique à celui utilisé pour les protéines natives.

### 4. Purification des mutants de Pa-aIF2 $\gamma$ .

Pour les mutants  $\gamma_{\Delta L1}$ ,  $\gamma_{\Delta L2}$ ,  $\gamma_{\Delta L3}$ ,  $\gamma_{\Delta L4}$ ,  $\gamma_{P178A}$ ,  $\gamma_{N186A}$  et  $\gamma_{C63H}$ , le même protocole de purification que pour aIF2 $\gamma$  native a été utilisé.

Pour les trois autres mutants,  $\gamma_{C60A}$ ,  $\gamma_{C72A}$  et  $\gamma_{\Delta Zn}$ , un très faible taux de surproduction a été obtenu dans les cultures d'un litre. Un meilleur taux de surproduction est obtenu sur des cultures de faible volume. Les purifications ont donc été réalisées à partir de cultures de 30 mL. Du fait de la très faible quantité d'aIF2 $\gamma$  présente dans les extraits protéiques, le protocole de purification classique n'a pas pu être appliqué pour ces mutants. En effet, après passage sur les quatre colonnes successives, la quantité de protéine purifiée était trop faible pour permettre une étude biochimique de ces mutants. Un autre protocole de purification, plus spécifique, a donc été envisagé. Il s'agit de purifier la protéine sur colonne d'affinité. Comme aucune colonne n'a pu retenir aIF2 $\gamma$  jusqu'à présent, une colonne contenant une résine TALON Superflow (CLONTECH Laboratories) a été utilisée (**Figure 6.6**).



**Figure 6.6 :** Principe de purification par chromatographie d'affinité sur colonne avec résine TALON.

Malheureusement, la protéine Pa-aIF2 $\gamma$  étiquetée par 6 His en N-terminal n'est pas retenue sur la colonne d'affinité de type TALON (Clontech). Par contre, l'hétérodimère  $\alpha\gamma$  où

$\alpha$  porte une queue polyhistidine en N-terminal est retenu sur une telle colonne d'affinité. Ce choix de purification a été validé par les essais préliminaires d'assemblage avec la sous-unité  $\gamma$  réalisés sur les extraits protéiques bruts obtenus après chauffage à 80°C des lysats cellulaires. Ces premiers essais semblent indiquer que les mutants du domaine de fixation du zinc n'interviennent pas dans la liaison à  $\alpha$ . Ces premiers tests ont donc permis de légitimer le choix de purification de l'hétérodimère  $\alpha\gamma$ .

La première étape a donc consisté à purifier la protéine aIF2 $\alpha$ -6xHis. Pour cela, le culot correspondant à 500 mL de culture est repris dans 30 mL de tampon B<sub>1</sub> (500 mM NaCl, 10 mM Tris pH 8.0, 3 mM 2-mercaptoethanol, 0.1 mM PMSF). Les cellules bactériennes sont alors lysées par sonication aux ultrasons pendant 6 min. (50% du temps actif) à 4°C. Après centrifugation (15300 rpm, 10 minutes, 4°C), le surnageant est chauffé à 80°C pendant 10 minutes. Après une nouvelle centrifugation (15300 rpm, 10 minutes, 4°C), le surnageant est directement chargé à 1 mL.min<sup>-1</sup> sur une colonne TALON (0,5 x 5 cm) équilibrée en tampon B<sub>1</sub>. La colonne est alors rincée en tampon B<sub>2</sub> (500 mM NaCl, 10 mM Tris pH 8.0, 3 mM 2-mercaptoethanol, 0.1 mM PMSF, 10 mM imidazole) sur 5 volumes de colonne. Puis la protéine est éluée en tampon B<sub>3</sub> (500 mM NaCl, 10 mM Tris pH 8.0, 3 mM 2-mercaptoethanol, 0.1 mM PMSF, 125 mM imidazole). Les fractions du pic d'éluion sont rassemblées et dialysées pendant 15 heures contre du tampon B<sub>1</sub>. Le dialysat est ensuite chargé à 2 mL.min<sup>-1</sup> sur une colonne Q-Hiload (1,1 x 4 cm) équilibrée en tampon B<sub>4</sub> (10 mM Tris pH 8.0, 3 mM 2-mercaptoethanol, 0.1 mM PMSF). Cette étape supplémentaire permet d'accrocher les acides nucléiques sur la colonne échangeuse d'anions tandis qu' $\alpha$ -6xHis passe au travers de la colonne. Les fractions correspondant à  $\alpha$  sont alors rassemblées. La colonne TALON est régénérée par rinçage (10 volumes de colonne) avec du tampon 20 mM MES pH 5.0 à un débit de 1 mL.min<sup>-1</sup>.

Pour la purification des hétérodimères  $\alpha\gamma_{\text{mutant}}$ , le culot des mutants de  $\gamma$  correspondant aux 30 mL de culture est repris dans 30 mL de tampon B<sub>1</sub>. La suspension bactérienne subit ensuite les mêmes traitements que dans le cas de la protéine  $\gamma$  native (lyse par sonication aux ultrasons, chauffage à 80°C, précipitation au sulfate de streptomycine puis au sulfate d'ammonium et dialyse contre du tampon B<sub>1</sub>). Avant injection sur la colonne d'affinité, un gel SDS-PAGE est réalisé pour vérifier la surproduction de la protéine  $\gamma$  mutante et estimer sa concentration. Au vu de cette quantification approximative, aIF2 $\alpha$ -6xHis est ajouté de manière à rester en défaut par rapport à  $\gamma$ . Le mélange  $\alpha\gamma$  est alors purifié sur la résine TALON suivant le même protocole que pour la protéine  $\alpha$  seule. L'excès de  $\gamma$  n'étant pas complexé à  $\alpha$ , il n'est pas retenu sur la colonne. Le pic d'éluion correspondant à l'hétérodimère est récolté et concentré par centrifugation sur membrane *Vivaspin 0.5 mL, 5000 MWCO*. Malgré la faible surproduction de ces mutants d'aIF2 $\gamma$ , une quantité suffisante

d'hétérodimère  $\alpha\gamma$  est obtenue dans chaque cas pour réaliser les tests biochimiques, notamment les tests d'assemblage avec la sous-unité  $\beta$ .

## **D. Cartographie de la zone d'interaction entre aIF2 $\alpha$ et aIF2 $\gamma$ .**

### **1. Obtention et purification des formes tronquées de Pa-aIF2 $\alpha$ .**

L'ADN codant pour le domaine 1 de la sous-unité  $\alpha$  ( $\alpha_{D1}$ ) a été obtenu par mutagenèse dirigée par la méthode *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene en introduisant un codon stop à la place du codon correspondant au résidu Gln88 de la protéine native. L'ADN codant pour la forme correspondant aux domaines 1 et 2 de la sous-unité  $\alpha$  ( $\alpha_{D1-2}$ ) a été obtenu selon le même protocole en introduisant un codon stop à la place du codon correspondant au résidu Val173 de la protéine native. Les produits de PCR obtenus ont ensuite été digérés par *DpnI* puis utilisés pour transformer des souches XL1 Blue d'*E. coli*.

Les ADN codant pour le domaine 3 ( $\alpha_{D3}$ ) et pour la forme correspondant aux domaines 2 et 3 ( $\alpha_{D2-3}$ ) de la sous-unité  $\alpha$  ont été obtenus par PCR en introduisant un codon de démarrage au niveau des résidus Val173 et Gln88 respectivement. Après amplification du gène muté, les produits de PCR ont été digérés successivement par *XbaI* et *SacII* puis reclés dans le pET3alpa digéré par les mêmes enzymes.

aIF2 $\alpha_{D1}$  et aIF2 $\alpha_{D3}$  sont purifiées selon le même protocole qu'aIF2 $\alpha$  native. Le tampon A<sub>4</sub> de la colonne S-Sepharose est cependant remplacé par du tampon A<sub>2</sub> (10 mM MOPS pH 6.7, 10 mM 2-mercaptoethanol, 100 mM NaCl). Les protéines sont ensuite éluées par gradient de 100 à 1000 mM NaCl en 60 minutes. Le pic d'éluion correspondant à la protéine (obtenu vers 450 mM NaCl) est rassemblé et la purification est ensuite poursuivie de manière identique à celle d'aIF2 $\alpha$ .

aIF2 $\alpha_{D1-2}$  et aIF2 $\alpha_{D2-3}$  sont purifiées selon le même protocole. Cependant, du fait de leur pI acide (6.5), la S-Sepharose est remplacée par une colonne Q-Hiload équilibrée en tampon A<sub>2</sub>.

### **2. Obtention et purification des formes tronquées de Pa-aIF2 $\beta$ .**

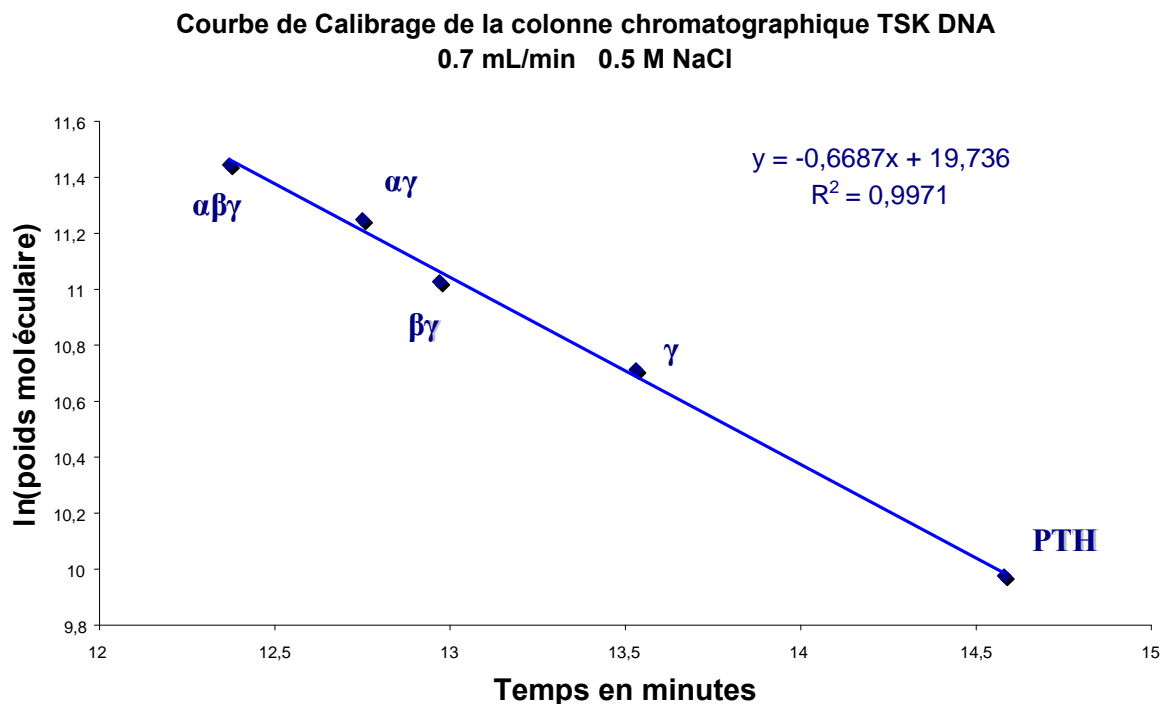
Les ADN codant pour le domaine 1 ( $\beta_1$ ) et pour le domaine 2 ( $\beta_2$ ) d'aIF2 $\beta$  ont été obtenus par PCR en introduisant respectivement un codon stop et un codon de démarrage à la place du codon correspondant au résidu Pro93 de la protéine native. Après amplification du

gène muté, les produits de PCR ont été digérés successivement par *XbaI* pour  $\beta_1$  (respectivement *NdeI* pour  $\beta_2$ ) et *SacII* pour puis reclésés dans le pET3alpha digéré par les mêmes enzymes.

Les protéines  $\beta_1$  et  $\beta_2$  sont purifiées suivant le même protocole qu'aIF2 $\beta$ .

### 3. Tests d'assemblage par chromatographie.

Les tests d'assemblage sont réalisés sur une colonne TSK3000SWXL (Tosohaas, Japon, 7.8 x 3000 mm<sup>2</sup>) équilibrée en tampon 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol, 500 mM NaCl. Pour chaque assemblage testé, les protéines purifiées correspondantes, natives ou mutées, sont mélangées (à une concentration de l'ordre de 5-10  $\mu$ M) de manière à avoir toujours un large excès d'aIF2 $\alpha$  et/ou aIF2 $\beta$  par rapport à la sous-unité  $\gamma$ . L'éluion est réalisée à 0.7 mL.min<sup>-1</sup>. L'attribution des différents pics chromatographiques pour les assemblages mutés est réalisée grâce à un calibrage préalable de la colonne avec les monomères et les assemblages natifs (**Figure 6.7**).



**Figure 6.7 :** Courbe d'étalonnage de la colonne TSK DNA (Tosohaas, Japon) pour les tests d'assemblage avec les mutants de la sous-unité  $\gamma$  par chromatographie d'exclusion.

### 4. Dichroïsme Circulaire.

Une fraction des solutions protéiques d'aIF2 $\gamma$  et d'aIF2 $\gamma_{\Delta L1}$  a été dialysée pendant 24 heures contre du tampon 10 mM phosphate de potassium pH 6.7, 100 mM NaCl pour

éliminer les nucléotides et le 2-mercaptoéthanol. La concentration précise de ces deux échantillons est ensuite déterminée en mesurant leur concentration en zinc par spectroscopie d'absorption atomique à l'aide d'un spectromètre de flamme *SpectrAA 220* (VARIAN, Palo Alto, Californie) selon la méthode décrite par Mayaux *et al.*, 1982. La sous-unité  $\gamma$  contient en effet un atome de zinc, de même que la sous-unité  $\beta$ . Brièvement, une gamme étalon de zinc (1.25 ; 2.5 ; 5 ; 10 ; 20  $\mu\text{M}$ ) est réalisée à partir d'une solution standard de zinc à 305.8 mM (Merck) diluée dans du tampon d'injection 10 mM phosphate de potassium pH 6.7, 10 mM 2-mercaptoéthanol, 100 mM NaCl, 0.1 mM EDTA. L'absorbance de chacune de ces solutions est mesurée à 213.9 nm. Une courbe d'étalonnage est ainsi réalisée et permet de mesurer la concentration précise en zinc des deux solutions protéiques après injection de 100  $\mu\text{L}$  d'une solution d'aIF2 $\gamma$  ou d'aIF2 $\gamma_{\Delta\text{L1}}$  diluée dans du tampon d'injection. La flamme du spectromètre provient de la combustion d'un mélange air-acétylène (débits respectifs : 13.5 et 1.5  $\text{L}\cdot\text{min}^{-1}$ ) et les échantillons sont injectés à un débit de 5  $\text{mL}\cdot\text{min}^{-1}$ .

Les mesures de dichroïsme circulaire sur aIF2 $\gamma$  et aIF2 $\gamma_{\Delta\text{L1}}$  ont été réalisées à l'Institut Curie d'Orsay sur un spectropolarimètre JASCO J-710 (2 nm de pas, vitesse de balayage : 50 nm/min) dans des cuves de quartz de 1 mm de parcours optique. Les concentrations en protéine utilisées pour l'enregistrement des spectres sont de l'ordre de 2.5  $\mu\text{M}$ . Les solutions protéiques récupérées après dialyse sont diluées dans le tampon de dialyse. Les spectres ont été enregistrés à 20°C entre 190 et 250 nm. Pour chaque spectre, 3 scans ont été enregistrés puis moyennés pour donner le spectre final. Des expériences de dénaturation thermique ont également été réalisées sur les deux protéines. Ces mesures ont été effectuées à 220 nm avec une montée en température de 40 à 97°C à 1°C.min<sup>-1</sup>.

### **5. Expériences de retard de migration sur gel.**

Les échantillons de 20  $\mu\text{L}$  sont préparés dans un tampon 50 mM Tris Borate EDTA (TBE) pH 8.4, 200 mM NaCl, 10 mM 2-mercaptoéthanol, 1 mM MgCl<sub>2</sub>, 1 mM GDPNP. Différentes concentrations de protéines (2.5 à 80  $\mu\text{M}$  selon les assemblages d'aIF2 testés) sont incubées en présence de 1.8  $\mu\text{M}$  d'ARNt<sub>f</sub><sup>Met</sup> aminoacylé ou non. L'incubation est maintenue 15 minutes à température ambiante. Puis les échantillons sont directement déposés sur un gel non dénaturant de composition suivante : 5% polyacrylamide (mono/bis, 19:1), 50 mM TBE pH 8.4, 1 mM MgCl<sub>2</sub> monté dans le système *Protean II* de Biorad. La migration se déroule à température ambiante pendant 45 minutes à 100V dans un tampon 50 mM TBE pH 8.4, 1 mM MgCl<sub>2</sub>. Après migration, les gels sont agités 5 à 10 minutes dans un bain de BET. Puis ils sont révélés sous éclairage UV.

## **II. ÉTUDE STRUCTURALE DE LA SOUS-UNITÉ $\alpha$ d'aIF2 de *P. abyssi*.**

### **A. Etude structurale de la forme tronquée Pa-aIF2 $\alpha_{D2-3}$ .**

#### **1. Mutation de Pa-aIF2 $\alpha_{D2-3}$ pour obtenir une protéine sélénométhionylable.**

La protéine  $\alpha_{D2-3}$  ne contenant pas de méthionine, son gène a été modifié par mutagenèse dirigée pour pouvoir incorporer de la sélénométhionine dans la protéine. Deux codons ATG codant pour la méthionine ont été introduits dans la séquence codant pour aIF2 $\alpha_{D2-3}$  : un codon ATG remplace le codon codant pour le résidu Leu106 et l'autre remplace le codon du résidu Ile184. Le double mutant a été construit en une seule étape de PCR par la méthode *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene en utilisant les deux amorces décrites dans le **Tableau 6.7** suivant :

<b>Amorces</b>	<b>Séquence</b>
Amorce 1	5' GCTCAGAAGGCAGAGAATATGCTCAAGCTCGCGGCCGAA 3'
Amorce 2	5' ATTTGGCTTTGGAACCGTCATCTCGAACTCCGCATCTAT 3'

**Tableau 6.7 :** Amorces utilisées pour synthétiser le double mutant d'aIF2 $\alpha_{D2-3}$  de *P.abyssi*.

L'amorce 1 permet d'introduire le codon Met en position 106 sur l'un des brins et l'amorce 2 permet d'introduire le codon Met en position 184 sur le brin complémentaire. Après amplification, l'ADN natif méthylé est digéré par *DpnI* avant transformation dans les cellules XL1 Blue d'*E. coli*. L'obtention du double mutant  $\alpha_{D2-3}$ L106M-I184M est vérifiée par séquençage.

#### **2. Production et purification de la protéine Pa-aIF2 $\alpha_{D2-3}$ sélénométhionylée.**

Le plasmide muté pET3a $\alpha_{D2-3}$ (L106M-I184M)lpa obtenu est utilisé pour transformer la souche d'*Escherichia coli* B834 (DE3) auxotrophe pour la méthionine. Les cellules sont mises en culture dans 60 mL de milieu LB en présence de 100  $\mu$ g/mL d'ampicilline. Cette préculture est incubée à 37°C durant 15 heures. Puis les cellules sont récupérées par centrifugation (8000g, 4°C, 20 minutes). Le culot est rincé deux fois en conditions stériles avec 60 mL d'eau stérile puis il est repris dans 30 mL de milieu minimal M9. 2.5 mL de cette suspension cellulaire sont alors introduits dans 1 L de milieu M9 contenant du glucose 0.2%, de l'ampicilline à 50  $\mu$ g/mL, de la kanamycine à 50  $\mu$ g/mL et 40 mg de chacun des acides aminés à l'exception de la méthionine. 15 mg de sélénométhionine (SIGMA) sont alors introduits dans la culture et celle-ci est incubée à 37°C jusqu'à ce que la DO<sub>650 nm</sub> atteigne 0.8-

1.0. La surproduction de la protéine est alors induite par ajout d'IPTG à une concentration finale de 0.1 mM et l'induction est poursuivie à température ambiante pendant 6 heures.

La protéine sélénométhionylée est purifiée suivant le même protocole que la protéine native.

### **3. Cristallisation de la protéine Pa-aIF2 $\alpha_{D2-3}$ et acquisition des données cristallographiques.**

La protéine  $\alpha_{D2-3}$  a été cristallisée par la méthode de diffusion de vapeur en goutte suspendue à l'aide d'un robot cristallographique *Genesis RSP100* (TECAN). Une solution protéique à 7 mg/mL ou 14 mg/mL dans un tampon 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol, 100 mM NaCl, 0.02% azide de sodium a été utilisée pour la recherche des conditions de cristallisation à l'aide des matrices d'échantillonnage commerciales fournies par Hampton Research. Les cristaux d'aIF2 $\alpha_{D2-3}$  ont été obtenus à 24°C au bout d'une semaine en utilisant la solution précipitante 5% isopropanol, 2.0 M sulfate d'ammonium (condition n°5 de la matrice d'échantillonnage *Crystal Screen 2*). Les cristaux de protéine sélénométhionylée ont été obtenus dans les mêmes conditions. Avant acquisition des données, les cristaux ont été trempés dans une solution 2.0 M sulfate d'ammonium contenant 25% de glycérol jouant le rôle de cryoprotectant. Ces cristaux ont été immédiatement refroidis soit dans l'éthane liquide, soit dans un flux d'azote. Les cristaux d'aIF2 $\alpha_{D2-3}$  appartiennent au groupe d'espace P2<sub>1</sub>2<sub>1</sub>2.

Les données de diffraction relatives aux cristaux cryoprotégés ont été collectées à 100K à l'ESRF (Grenoble) sur les lignes ID14-eh1 (cristaux natifs) et ID23-eh1 (cristaux sélénométhionylés). Un jeu de données à 2.26 Å a été enregistré pour les cristaux natifs sur un détecteur de type *Quantum 4R CCD* (ADSC, Poway, Californie). Pour les cristaux sélénométhionylés, deux jeux de données à 2.8 Å ont été enregistrés grâce à un détecteur *MarMosaic 225 CCD* (Mar Research, Hambourg, Allemagne): un jeu à la longueur d'onde correspondant au maximum d'absorption du sélénium, soit 0.979 nm, et un jeu à la longueur d'onde correspondant à l'inflexion du pic d'absorption du Se, soit 0.9756 nm.

### **4. Résolution et affinement de la structure de la protéine Pa-aIF2 $\alpha_{D2-3}$ .**

Les clichés de diffraction obtenus pour les cristaux natifs et sélénométhionylés ont été traités à l'aide du programme *MOSFLM* (Leslie, 1990). Les données ont ensuite été analysées avec les programmes développés dans le cadre du CCP4 (Collaborative Computational Project N°4, 1994).

La structure d'eIF2 $\alpha_{D2-3}$  a été résolue à l'aide du programme *SOLVE* (Terwilliger and Berendzen, 1999). Les deux jeux de données obtenus pour les cristaux sélénométhionylés ont été utilisés comme paramètres d'entrée du programme *SOLVE*. Les positions des deux atomes de sélénium attendus ont été trouvées par le programme. Puis leurs positions ont été affinées. Ces deux sites ont ensuite été utilisés pour trouver les phases du jeu de réflexions et tracer une carte de densité initiale. Cette carte est alors améliorée par lissage de solvant à l'aide du programme *RESOLVE* (Terwilliger, 1999). La carte ainsi obtenue présente une densité électronique suffisamment organisée pour positionner le modèle structural correspondant au domaine 2 de la protéine eIF2 $\alpha$  humaine (Nonato *et al.*, 2002) et identifier les éléments de structure relatifs au domaine 3. La construction du modèle final est effectuée à l'aide du programme *O* (Jones *et al.*, 1991) au fur et à mesure des tours d'affinements. L'affinement du modèle a été réalisé par différents cycles de minimisation et de dynamique moléculaires en utilisant le programme *CNS* (Brunger *et al.*, 1998). Le modèle final ainsi construit présente un facteur R de 23.3% et un facteur R<sub>libre</sub> de 28%.

## **B. Etude structurale de la protéine entière Pa-eIF2 $\alpha$ .**

### **1. Cristallisation de la protéine Pa-eIF2 $\alpha$ et acquisition des données cristallographiques.**

La protéine  $\alpha$  a été cristallisée manuellement par la méthode de diffusion de vapeur en goutte suspendue. Une solution protéique à 15 mg/mL dans un tampon 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol, 500 mM NaCl, 0.02% azide de sodium a été utilisée pour la recherche des conditions de cristallisation. Les cristaux d'eIF2 $\alpha$  ont été obtenus à 24°C au bout de deux jours en utilisant la solution précipitante 0.6 M sulfate d'ammonium, 0.1 M NaAc pH 5.0. Pour l'acquisition des données, les cristaux ont été trempés dans une solution 0.5 M sulfate d'ammonium, 0.83 M NaAc pH 5.0 contenant 40% d'éthylène glycol cryoprotecteur et immédiatement refroidis dans l'éthane liquide.

Les données de diffraction relatives à ces cristaux ont été collectées à 100K à l'ESRF (Grenoble) sur la ligne ID14-eh1. Un jeu de données à 3.37 Å a été enregistré pour ces cristaux qui appartiennent au groupe d'espace P3<sub>1</sub>21 sur un détecteur de type *Quantum 4R CCD* (ADSC, Poway, Californie).

### **2. Résolution et affinement de la structure de la protéine eIF2 $\alpha$ .**

Les clichés de diffraction obtenus pour les cristaux d' $\alpha$  ont été traités à l'aide du programme *MOSFLM* (Leslie, 1990). Les données ont ensuite été analysées avec les

programmes développés dans le cadre du CCP4 (Collaborative Computational Project N°4, 1994).

La structure de Pa-aIF2 $\alpha$  a ensuite été résolue par remplacement moléculaire à l'aide du programme *PHASER* (Storoni *et al.*, 2004). La structure de Pa-aIF2 $\alpha_{D2-3}$  précédemment résolue a servi de modèle pour les domaines 2 et 3 de la protéine tandis que le domaine 1 d'eIF2 $\alpha$  de levure (Dhaliwal and Hoffman, 2003) a été utilisé pour modéliser le domaine 1 de Pa-aIF2 $\alpha$ . Le domaine 3 a été cherché en premier à partir du modèle du domaine 3 de la structure d'aIF2 $\alpha_{D2-3}$  de *P. abyssi*. Un Zscore de 17.2 a été obtenu pour ce domaine. A partir de cette solution, le domaine 2 a été cherché à partir du modèle du domaine 2 de la structure d'aIF2 $\alpha_{D2-3}$ . Une solution unique avec un Zscore de 14.9 a été obtenue. Pour finir, le domaine 1 a été cherché à partir du modèle du domaine 1 d'eIF2 $\alpha$  de levure. Un Zscore de 10.9 a été obtenu avec ce modèle. Ces trois solutions uniques ont permis de construire un modèle initial d'aIF2 $\alpha$  ainsi qu'une première carte de densité. Les résidus du polypeptide correspondant au domaine 1 d' $\alpha$  ont ensuite été changés en alanine à l'exception des résidus communs au domaine 1 d'eIF2 $\alpha$  de levure et au domaine 1 d'aIF2 $\alpha$  de *P. abyssi*. Les résidus du domaine 1 d'aIF2 $\alpha$  de *P. abyssi* ont ensuite été remplacés au fur et à mesure des tours d'affinement. La construction du modèle final a été effectuée à l'aide du programme *O* (Jones *et al.*, 1991) et le modèle a été affiné par différents cycles de minimisation et de dynamique moléculaires à l'aide du programme *CNS* (Brunger *et al.*, 1998) en utilisant seulement les données comprises entre 12 et 3.37 Å et dont la valeur de l'amplitude est supérieure à deux fois la valeur de l'écart type ( $>2\sigma$ ). Le modèle final ainsi construit présente un facteur R de 25% et un facteur  $R_{\text{libre}}$  de 29.2%.

### **3. Analyse de la conservation de séquence.**

66 séquences d'e/aIF2 $\alpha$  ont été identifiées en utilisant le programme *BLAST* de la base de données PubMed ([www.ncbi.nlm.nih.gov/blast](http://www.ncbi.nlm.nih.gov/blast)). Ces 66 séquences regroupent 43 séquences de protéines  $\alpha$  eucaryotiques et 23 séquences de protéines  $\alpha$  d'Archées. L'alignement de ces séquences a été réalisé à l'aide du programme *Clustal X* (Thompson *et al.*, 1997). La conservation de chaque position a été étudiée à l'aide du programme *Plotsimilarity* du package GCG disponible sur le site d'Infobiogen ([www.infobiogen.fr](http://www.infobiogen.fr)), en utilisant une fenêtre de lecture de 5 résidus.

### **III. ÉTUDE FONCTIONNELLE ET STRUCTURALE DE L'HÉTÉRODIMÈRE $\alpha\text{IF}2\alpha\gamma$ .**

#### **A. L'hétérodimère $\alpha\text{IF}2\alpha_{D3}\gamma$ de *P. abyssi*.**

##### **1. Cristallisation de l'hétérodimère $\alpha\text{IF}2\alpha_{D3}\gamma$ de *P. abyssi* et acquisition des données cristallographiques.**

L'hétérodimère  $\alpha_{D3}\gamma$  a été cristallisé manuellement par la méthode de diffusion de vapeur en goutte suspendue (gouttes de 4  $\mu\text{L}$  : 2  $\mu\text{L}$  de solution protéique + 2  $\mu\text{L}$  de solution réservoir). Une solution protéique à 6.6 mg/mL dans un tampon 10 mM MOPS pH 6.7, 10 mM 2-mercaptoéthanol, 200 mM NaCl, 0.02% azide de sodium, 1 mM GDPNP, 1 mM  $\text{MgCl}_2$  a été utilisée pour la recherche des conditions de cristallisation à l'aide des matrices d'échantillonnage commerciales fournies par Hampton Research. Les cristaux d' $\alpha\text{IF}2\alpha_{D3}\gamma$  ont été obtenus à 24°C au bout d'une journée en utilisant la condition n°36 de la matrice *Crystal Screen 2* (4.3 M NaCl, 0.1 M HEPES pH 7.5). Pour l'acquisition des données, les cristaux d' $\alpha\text{IF}2\alpha_{D3}\gamma$  ont été trempés dans une solution contenant 4.22 M NaCl, 0.1 M HEPES pH 7.5 et 10% éthylène glycol (cryoprotectant) et immédiatement refroidis dans l'éthane liquide. Les données de diffraction relatives à ces cristaux ont été collectées à l'ESRF (Grenoble) sur la ligne ID14-eh3. Un jeu de données à 5.6 Å a été enregistré pour ces cristaux qui appartiennent au groupe d'espace  $P3_221$  sur un détecteur de type *Quantum 4R CCD* (ADSC, Poway, Californie).

Compte tenu du faible pouvoir diffractant de ces cristaux, nous avons tenté d'améliorer la qualité des cristaux en optimisant les conditions de cristallisation à l'aide des matrices d'échantillonnages *Additives* et *Detergents* d'Hampton Research. Les différents composés présents dans ces solutions commerciales ont été introduits à une dilution de 1/10<sup>ème</sup> dans les gouttes de cristallisation de l'hétérodimère  $\alpha_{D3}\gamma$  en présence de la solution précipitante 4.3 M NaCl, 0.1 M HEPES pH 7.5. Les différents cristaux obtenus pour chaque additif ont subi le même traitement de cryoprotection que les cristaux initiaux. Leur pouvoir diffractant a été testé à l'ESRF (Grenoble) sur la ligne ID14-eh3. Le meilleur jeu de données a été obtenu avec l'additif HEGA-9 présent à la concentration de 39 mM dans la goutte de cristallisation. Un jeu à 4.7 Å a ainsi été enregistré.

##### **2. Tentative de résolution de la structure de l'hétérodimère $\alpha\text{IF}2\alpha_{D3}\gamma$ de *P. abyssi*.**

Les clichés de diffraction obtenus pour les cristaux d' $\alpha_{D3}\gamma$  ont été traités à l'aide du programme *MOSFLM* (Leslie, 1990). Les données ont ensuite été analysées avec les

programmes développés dans le cadre du CCP4 (Collaborative Computational Project N°4, 1994).

La structure d' $\alpha_{D3}\gamma$  a ensuite été résolue par remplacement moléculaire à l'aide du programme *PHASER* (Storoni *et al.*, 2004). Les structures des protéines  $\alpha_{D2-3}$  et  $\alpha_{D2-3}\gamma$  de *P. abyssi* ont servi de modèles pour résoudre cette structure.

Le domaine 1 d' $\alpha_{D2-3}\gamma$  a été cherché en premier à partir du modèle du domaine 1 de la structure d' $\alpha_{D2-3}\gamma$  de *P. abyssi*. Un Zscore de 9.9 a été obtenu pour ce domaine. A partir de cette solution, le domaine 2 d' $\alpha_{D2-3}\gamma$  a été cherché à partir du modèle du domaine 2 de la structure d' $\alpha_{D2-3}\gamma$  de *P. abyssi*. Une solution unique avec un Zscore de 14.5 a été obtenue. Puis, le domaine 3 a été cherché à partir du modèle du domaine 3 d' $\alpha_{D2-3}\gamma$  de *P. abyssi*. Un Zscore de 19.5 a été obtenu avec ce modèle. Pour finir, le domaine 3 d' $\alpha_{D2-3}$  a été cherché à partir du modèle du domaine 3 d' $\alpha_{D2-3}\gamma$  de *P. abyssi*. Un Zscore de 6.3 a été obtenu. Ces quatre solutions uniques ont permis de construire une première carte de densité qui a été améliorée par lissage de solvant à l'aide du programme *DM* (Cowtan and Main, 1993) du package CCP4 (Collaborative Computational Project N°4, 1994). Le modèle a alors été affiné par une minimisation d'énergie en considérant les trois domaines de  $\gamma$  et le domaine 3 d' $\alpha$  comme des ensembles rigides à l'aide du programme *CNS* (Brunger *et al.*, 1998). Un facteur R de 40 % et un facteur  $R_{\text{libre}}$  de 42% ont ainsi été obtenus. Les tentatives de remplacement manuel du modèle dans la densité électronique et les différents tours d'affinement réalisés n'ont pas permis d'améliorer ces paramètres du fait du peu de réflexions mesurées à 4.7 Å de résolution.

### **B. Production et purification des différentes sous-unités du facteur hétérotrimérique $\alpha_{IF2}$ de *S. solfataricus*.**

Les ADN codant pour chacune des 3 sous-unités du facteur hétérotrimérique Ss- $\alpha_{IF2}$  ont été clonés dans le vecteur pET3 $\alpha$  entre les sites de restriction de *NdeI* et *XhoI* du fragment « lpa » additionnel, donnant ainsi naissance aux vecteurs pET3 $\alpha$ sslpa, pET3 $\beta$ sslpa et pET3 $\gamma$ sslpa.

Ces vecteurs ont ensuite été utilisés pour transformer des cellules Rosetta pLacIRARE d'*Escherichia coli*. Ces cellules contiennent le plasmide pLacIRARE qui permet la surexpression d'ARNts capables de lire les codons rares et la surexpression de lacI. Ces cellules transformées sont alors mises en culture à 37°C dans 500 mL de milieu 2xTY contenant 50  $\mu$ g/mL d'ampicilline et 34  $\mu$ g/mL de chloramphénicol. L'expression des protéines d' $\alpha_{IF2}$  est alors induite par ajout d'IPTG à une concentration finale de 1 mM

lorsque la densité optique de la culture à 650 nm atteint 1,6. L'induction est poursuivie durant 6 heures à température ambiante. Puis les cultures sont centrifugées à 8000 g pendant 30 minutes et les culots sont congelés et conservés à -20°C jusqu'à utilisation.

Les ADNs codant pour le domaine 3 ( $\alpha_{D3}$ ) et pour la forme correspondant aux domaines 2 et 3 ( $\alpha_{D2-3}$ ) de la sous-unité  $\alpha$  ont été obtenus par PCR en introduisant un codon de démarrage au niveau des résidus Glu174 et Asp88 respectivement. Après amplification du gène muté, les produits de PCR ont été digérés successivement par *NdeI* et *XhoI* puis reclésés dans le pET3alpa digéré par les mêmes enzymes.

### **1. Purification de Ss-aIF2.**

Les culots de cellules correspondant à 500 mL de culture de chaque sous-unité sont resuspendus dans du tampon C<sub>1</sub> (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM benzamidine, 0.1 mM PMSF, 100 mM NaCl). Les différentes suspensions sont ensuite mélangées dans un volume final d'environ 100 mL. Du GTP (à la concentration finale de 100  $\mu$ M) et du MgCl<sub>2</sub> (1 mM) sont ajoutés pour stabiliser la protéine Ss-aIF2 $\gamma$  et éviter en particulier la protéolyse au niveau de la boucle GKT impliquée dans la liaison des nucléotides G. Les cellules bactériennes sont alors lysées par sonication aux ultrasons pendant 8 min. (50% du temps actif) à 4°C. Après centrifugation (15300 rpm, 10 minutes, 4°C), le surnageant est chauffé à 65°C pendant 10 minutes pour précipiter les protéines bactériennes thermosensibles. Ces dernières sont alors séparées des protéines thermostables par centrifugation (15300 rpm, 10 minutes, 4°C).

Le surnageant est directement chargé à 2.5 mL.min<sup>-1</sup> sur une colonne échangeuse de cations S-Hiload (1,6 x 20 cm) équilibrée en tampon C<sub>1</sub>. L'hétérotrimère est ensuite élué avec le tampon C<sub>2</sub> (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM benzamidine, 0.1 mM PMSF, 400 mM NaCl). Les fractions d'élué contenant Ss-aIF2 sont rassemblées et concentrées par centrifugation sur membrane *Vivaspin 20 mL, 30 000 MWCO* jusqu'à un volume inférieur à 5 mL.

Pour finir, l'échantillon est chargé à 0.4 mL.min<sup>-1</sup> sur un tamis moléculaire Superdex 200 (1,6 x 60 cm) équilibré en tampon C<sub>3</sub> (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM benzamidine, 0.1 mM PMSF, 200 mM NaCl). L'hétérotrimère sort sous forme d'un pic homogène. Les fractions correspondantes sont rassemblées et concentrées par centrifugation sur membrane *Amicon-Ultra 4 mL, 30 000 MWCO* jusqu'à atteindre une concentration de l'ordre de 8-10 mg/mL. De l'azide de sodium à 0,02% final est ajouté pour assurer la stabilité de la protéine. La solution protéique est

ensuite divisée en trois échantillons : un échantillon contenant en plus du GDPNP et du  $MgCl_2$  à une concentration finale de 1 mM, un échantillon avec GDP et  $MgCl_2$  (1 mM) et le dernier échantillon sans nucléotide guanine ni magnésium, dans le but de réaliser des essais de cristallisation de Ss-aIF2 en présence de l'un ou l'autre des nucléotides guanine ou sous forme apo. L'hétérotrimère est alors conservé à 4°C. En moyenne, 25 mg de protéine purifiée sont obtenus à partir de 500 mL de culture de chacune des sous-unités.

## **2. Purification des hétérodimères Ss-aIF2 $\alpha\gamma$ , $\alpha_{D3}\gamma$ et $\alpha_{D2-3}\gamma$ .**

Le même protocole de purification que celui de l'hétérotrimère natif a été appliqué pour purifier ces différents assemblages. La Superdex 200 (1,6 x 60 cm) a été remplacée par une Superdex 75 (1,6 x 60 cm).

## **3. Purification de Ss-aIF2 $\gamma$ .**

Le culot de cellules correspondant à 500 mL de culture est resuspendu dans environ 50 mL de tampon C<sub>4</sub> (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM benzamidine, 0.1 mM PMSF, 500 mM NaCl). Du GTP (100  $\mu$ M) et du  $MgCl_2$  (1 mM) sont ajoutés pour stabiliser la protéine. Les cellules bactériennes sont alors lysées par sonication aux ultrasons pendant 8 min. (50% du temps actif) à 4°C. Après centrifugation (15300 rpm, 10 minutes, 4°C), le surnageant est chauffé à 65°C pendant 10 minutes pour précipiter les protéines bactériennes thermosensibles. Ces dernières sont alors séparées des protéines thermostables par centrifugation (15300 rpm, 10 minutes, 4°C).

Le surnageant est directement chargé sur une colonne échangeuse d'anions Q-Hiload (1,6 x 20cm) équilibrée en tampon C<sub>4</sub> à 2.5 mL.min<sup>-1</sup>. La protéine n'est pas retenue sur cette colonne qui permet d'éliminer une grande partie des acides nucléiques. Les fractions contenant Ss-aIF2 $\gamma$  sont rassemblées, 1 mM GTP et 1 mM  $MgCl_2$  sont ajoutés, et l'échantillon est mis à dialyser 1 heure contre 1 L de tampon C<sub>5</sub> (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM benzamidine, 0.1 mM PMSF).

La seconde étape consiste en une colonne échangeuse de cations S-Hiload (1,6 x 20 cm). Le dialysat précédent est chargé à 2.5 mL.min<sup>-1</sup> sur cette colonne équilibrée en tampon C<sub>5</sub>. La protéine est éluée avec le tampon C<sub>2</sub> (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM benzamidine, 0.1 mM PMSF, 400 mM NaCl). Les fractions d'éluion contenant Ss-aIF2 $\gamma$  sont rassemblées, 1 mM GTP et 1 mM  $MgCl_2$  sont ajoutés, et l'échantillon est concentré par centrifugation sur membrane *Vivaspin 20 mL, 30 000 MWCO* jusqu'à un volume inférieur à 5 mL.

Pour finir, l'échantillon est chargé à  $0.4 \text{ mL} \cdot \text{min}^{-1}$  sur un tamis moléculaire Superdex 75 (1,6 x 60 cm) équilibré en tampon C<sub>3</sub>. Les fractions correspondant à Ss-aIF2 $\gamma$  sont rassemblées et concentrées par centrifugation sur membrane *Amicon-Ultra 4 mL, 30 000 MWCO* jusqu'à atteindre une concentration de l'ordre de 8-10 mg/mL. En moyenne, 5 mg de protéine purifiée sont obtenus à partir de 500 mL de culture.

#### **4. Obtention et purification des mutants ponctuels d'aIF2 $\alpha$ de *S. solfataricus*.**

Les mutants ponctuels R186A, R226A et Y227A d'aIF2 $\alpha$  ont été construits par mutagenèse dirigée par PCR par la méthode *QuikChange<sup>TM</sup> Site-Directed Mutagenesis* de Stratagene à partir du vecteur cible pET3a $\alpha$ sslpa. Les vecteurs mutés résultants ont été séquencés pour contrôler l'obtention de la mutation désirée. Les protéines correspondantes ont été produites et purifiées suivant le même protocole que la protéine native.

Avant de purifier les hétérodimères  $\alpha\gamma$  correspondants, un gel natif acide en conditions non dénaturantes a été réalisé pour tester la capacité de ces mutants à s'assembler avec la sous-unité  $\gamma$ . Aucun défaut d'association n'ayant été observé, les trois hétérodimères  $\alpha_{R186A}\gamma$ ,  $\alpha_{R226A}\gamma$  et  $\alpha_{Y227A}\gamma$  ont été purifiés selon le même protocole que pour l'hétérodimère natif.

#### **6. Détermination de la constante de dissociation des complexes aIF2:ARNt.**

Le même protocole que pour les protéines de *P. abyssi* a été utilisé. Cependant, une concentration de 2 nM en [<sup>35</sup>S]-Met-ARNt<sub>f</sub><sup>Met</sup> a été utilisée dans le test de protection contre la déacylation.

### **C. Etude structurale de l'hétérodimère aIF2 $\alpha\gamma$ de *S. solfataricus*.**

#### **1. Cristallisation de l'hétérodimère aIF2 $\alpha\gamma$ de *P. abyssi* et acquisition des données cristallographiques.**

L'hétérodimère  $\alpha\gamma$  a été cristallisé par la méthode de diffusion de vapeur en goutte suspendue à l'aide d'un robot cristallographique *Genesis RSP100* (TECAN). Une solution protéique à 9 mg/mL dans un tampon 10 mM HEPES pH 7.5, 10 mM 2-mercaptoéthanol, 0.1 mM EDTA, 0.1 mM PMSF, 0.1 mM benzamidine, 200 mM NaCl, 0.02% azide de sodium, 1 mM GDPNP, 1 mM MgCl<sub>2</sub> a été utilisée pour la recherche des conditions de cristallisation à l'aide des matrices d'échantillonnage commerciales fournies par Hampton Research. Les

cristaux d'aIF2 $\alpha\gamma$  ont été obtenus à 24°C au bout d'une semaine en utilisant les solutions précipitantes 10% PEG 6000, 0.1 M HEPES pH 7.0 et 10% PEG 6000, 0.1 M Bicine pH 9.0 (conditions n°10 et 12 de la matrice d'échantillonnage *Grid Screen PEG 6000*). De manière surprenante, ces cristaux se sont révélés peu stables dans l'éthane liquide après cryogénéisation. Ils sont donc directement montés dans les boucles lors de l'acquisition des données. Au préalable, ces cristaux sont trempés dans une solution contenant la solution précipitante et 20% d'éthylène glycol comme cryoprotectant. Ils sont alors immédiatement refroidis dans le flux d'azote du montage et l'enregistrement des données de diffraction est lancé. Les cristaux d'aIF2 $\alpha\gamma$  appartiennent au groupe d'espace P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>.

Les données de diffraction relatives aux cristaux d' $\alpha\gamma$  ont été collectées à 100K à l'ESRF (Grenoble) sur la ligne ID14-eh1. Un jeu de données à 3.0 Å a été enregistré sur un détecteur de type *Quantum 4R CCD* (ADSC, Poway, Californie).

## **2. Résolution et affinement de la structure de l'hétérodimère aIF2 $\alpha\gamma$ .**

Les clichés de diffraction obtenus pour les cristaux d' $\alpha\gamma$  ont été traités à l'aide du programme *XDS* (Kabsch, 1988).

La structure d'aIF2 $\alpha\gamma$  a ensuite été résolue par remplacement moléculaire à l'aide du programme *PHASER* (Storoni *et al.*, 2004). Les trois domaines séparés d'aIF2 $\gamma$  ont d'abord été cherchés. La structure d'aIF2 $\gamma$  de *P. abyssi* (Schmitt *et al.*, 2002) a servi de modèle pour ces domaines. Le domaine 1 a été cherché en premier à partir du modèle du domaine 1 de la structure d'aIF2 $\gamma$  de *P. abyssi*. Un Zscore de 7.6 a été obtenu pour ce domaine. A partir de cette solution, le domaine 2 a été cherché à partir du modèle du domaine 2 de la structure d'aIF2 $\gamma$  de *P. abyssi*. Une solution unique avec un Zscore de 11.6 a été obtenue. Pour finir, le domaine 3 a été cherché à partir du modèle du domaine 3 d'aIF2 $\gamma$  de *P. abyssi*. Un Zscore de 14.1 a été obtenu avec ce modèle. Pour la sous-unité  $\alpha$ , différents modèles structuraux ont ensuite été utilisés : aIF2 $\alpha_{D2-3}$  de *P. abyssi* précédemment résolue (modèle entier ou domaines séparés), aIF2 $\alpha$  de *P. abyssi* précédemment résolue (modèle entier ou domaines séparés), eIF2 $\alpha_{D1-2}$  de levure (Dhaliwal and Hoffman, 2003), eIF2 $\alpha_{D1-2}$  humain (Nonato *et al.*, 2002) et eIF2 $\alpha$  humain (Ito *et al.*, 2004). Aucune solution pour la sous-unité  $\alpha$  n'a pu être obtenue par remplacement moléculaire quel que soit le modèle structural utilisé.

La résolution de la structure du complexe entier a alors été obtenue comme suit. Les trois solutions uniques obtenues pour les trois domaines d'aIF2 $\gamma$  ont donc été utilisées pour construire un modèle initial d'aIF2 $\gamma$  ainsi qu'une première carte de densité. aIF2 $\gamma$  de *S. solfataricus* présente 56% d'identité de séquence avec aIF2 $\gamma$  de *P. abyssi*. Dans un premier

temps, les résidus divergents entre les deux protéines ont été changés en alanine sur le modèle structural obtenu par remplacement moléculaire. La carte de densité initiale a alors été améliorée par lissage de solvant à l'aide du programme *DM* (Cowtan and Main, 1993) du package CCP4 (Collaborative Computational Project N°4, 1994). Puis, le modèle a été affiné avec différents cycles de construction manuelle et de minimisation d'énergie à l'aide des programmes *O* (Jones *et al.*, 1991) et *CNS* (Brunger *et al.*, 1998) en utilisant seulement les données comprises entre 12 et 3.0 Å et dont la valeur de l'amplitude est supérieure à deux fois la valeur de l'écart type. L'affinement du modèle sur la protéine  $\gamma$  seule a permis d'obtenir une carte suffisamment organisée pour pouvoir positionner le domaine 3 d'aIF2 $\alpha$  de *P. abyssi* dans la densité électronique. Là encore, les résidus divergents entre aIF2 $\alpha$  de *S. solfataricus* et aIF2 $\alpha$  de *P. abyssi* ont été changés en alanine dans un premier temps. Puis les résidus corrects ont été réinsérés au fur et à mesure des tours d'affinements. Finalement, les domaines 2 puis 1 d'aIF2 $\alpha$  de *P. abyssi* ont pu être positionnés sans ambiguïté dans la densité électronique et le modèle a finalement été affiné pour obtenir un facteur R final de 23.2% et un facteur R<sub>libre</sub> de 29.7% à 3.0 Å de résolution.



# *ANNEXES*



Annexes

e/aIF2γ

Table with 2 columns: Species and protein sequence. Species include S. cerevisiae, S. pombe, H. sapiens, B. taurus, M. musculus, D. melanogaster, A. thaliana, A. fulgidus, M. thermoautotrophicum, P. horikoshii, M. jannaschii, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

EF1A-e/aEF1A

Table with 2 columns: Species and protein sequence. Species include T. thermophilus, E. coli, S. cerevisiae, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

e/aIF2γ

Table with 2 columns: Species and protein sequence. Species include S. cerevisiae, S. pombe, H. sapiens, B. taurus, M. musculus, D. melanogaster, A. thaliana, A. fulgidus, M. thermoautotrophicum, P. horikoshii, M. jannaschii, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

EF1A-e/aEF1A

Table with 2 columns: Species and protein sequence. Species include T. thermophilus, E. coli, S. cerevisiae, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

e/aIF2γ

Table with 2 columns: Species and protein sequence. Species include S. cerevisiae, S. pombe, H. sapiens, B. taurus, M. musculus, D. melanogaster, A. thaliana, A. fulgidus, M. thermoautotrophicum, P. horikoshii, M. jannaschii, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

EF1A-e/aEF1A

Table with 2 columns: Species and protein sequence. Species include T. thermophilus, E. coli, S. cerevisiae, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

e/aIF2γ

Table with 2 columns: Species and protein sequence. Species include S. cerevisiae, S. pombe, H. sapiens, B. taurus, M. musculus, D. melanogaster, A. thaliana, A. fulgidus, M. thermoautotrophicum, P. horikoshii, M. jannaschii, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

EF1A-e/aEF1A

Table with 2 columns: Species and protein sequence. Species include T. thermophilus, E. coli, S. cerevisiae, S. solfataricus, and P. abyssi. Sequences are aligned and color-coded.

**e/aIF2 $\gamma$**

S. cerevisiae KPIFSNIVSLFAEQNDLKFVPGSLIIVGTTKVDPTLCRADRLVQVVCAGKHLNIIY---TDIEINYFLRRLLGVKTDG----- 447  
 S. pombe QPIFSRIISLFAEHNDLKIAPVPGSLIIVGTTVDEPTLCRADRLVQVVLGSKGNLPEVY---TELEINYFLRRLLGKKS-GD----- 368  
 H. sapiens KPIFSKIVSLFAEHNDLQYAAPGSLIIVGTTKIDPTLCRADRMVQVLCVAVGALPEIF---TELEISYFLRRLLGVRTEGD----- 389  
 B. taurus KPIFSKIVSLFAEHNDLQYAAPGSLIIVGTTKIDPTLCRADRMVQVLCVAVGALPEIF---TELEISYFLRRLLGVRTEGD----- 389  
 M. musculus KPIFSKIVSLFAEHNDLQYAAPGSLIIVGTTKIDPTLCRADRMVQVLCVAVGALPEIF---TELEISYFLRRLLGVRTEGD----- 389  
 D. melanogaster RPIFSRIISLFAEQNELQYAVPGSLIIVGTTKIDPTLCRADRLVQVLCVAVGQLPEIY---QELEISYFLRRLLGVRTDGD----- 388  
 A. thaliana TPIYTRITSLYAEQNELQFVAPGSLIIVGTTSMDEPTLRGDRIVQVLCGEIGTLPDIY---VELEVSQQLTRLLGVKTKET----- 380  
 A. fulgidus NPIFTEVQSLIVASGRFVDEATPGSLVSIATKLDPTLTKSDALVENVVVCHPGNLEDVL---TSFTIMEVNLDERVVGL----- 353  
 M. thermoautotrophicum MSIHSTITGLVAGGEEMEEVPGGLVIVGTTLDPAITKADSLSGVAPEPTLPPVR---HSFTMETHLDERVVGT----- 337  
 P. horikoshii EPIITTEIVSLQAGGQFVEEAYPGSLVIVGTTKLDPEYLTGKDLMAENVVCKPGKLPVW---DSLRLVHLLDERVVGT----- 339  
 M. jannaschii KPIITKIVSLAAGNTILRKAHPGSLIIVGTTLDPEYLTGKSDALTSVVGSLPGLTPIR---EKITIRANLDRVVGT----- 339  
 S. solfataricus EPIITKISSIRFQDEEFKEAKPGSLVIVGTTLDPEYLTGKSDALTSVVGSLPGLTPIR---WNIRIKYNIIDERVVGA----- 342  
 P. abyssi EPIITTEIVSLQAGGQFVEEAYPGGLVIVGTTKLDPEYLTGKDLMAENVVCKPGKLPVW---DSLRLVHLLDERVVGT----- 339

**EF1A-e/aEF1A**

T. thermophilus ETRRTVVTGVEMHRKTLQEGIAQDNVIVLLRG---VSREEVERQVLAKPGSITPHT----KFEASVYVKK--EEGGRHTGFFSGYR 340  
 E. coli ETQKSTCTGVEMFRKLLDEGRACENVIVLLRG---IKREEIERQVLAKPGTIKPH-----KFESEVYISK--DEGGRHTPFPGYR 328  
 S. cerevisiae --VTTEVKIVEMHHEQLQEQGQDNVIVFNVK---VSVEKIEIRNVCSDAKNDPPKG---CASFNATVIVNHPGQISAGYSPVLDCHT 363  
 S. solfataricus --KVGEVRILETHHTKMDKABEQDNIVFNVRG---VEKKIIRKQDVVHPNPPETVA---DEFTARIIVVWHPTALANGYTPVIHVHT 351  
 P. abyssi KPIIQGEVKIVEMHHEPLQEQALQDNIVFNVRG---VSKNDIKRQDVAHPDKPETVVRTKDTFKAQIIVNHPATAITVGYSPVLHAHT 348

**e/aIF2 $\gamma$**

S. cerevisiae -QKQAKVRKLEPNEVIVMNIIGSTATGARVVAVKADMARIQITSPACTEIN-EKIALSRRIEKHWR-LICWATLKKGTTLEPIA----- 527  
 S. pombe -KNTTKVQKLAKNEVIVMNIIGSTISGGRVMMVKADMAKILITAPACTEIG-EKVALSRRIEKHWR-LICWAKVVEGKTLKV----- 446  
 H. sapiens -KKAARKVQKLSKNEVIVMNIIGSLSTGGRVSAVKADLGRIVITNIVCTEVG-EKIALSRRIEKHWR-LICWQIIRRGVTTIKPTVDDD-- 472  
 B. taurus -KKAARKVQKLSKNEVIVMNIIGSLSTGGRVSAVKADLGRIVITNIVCTEVG-EKIALSRRIEKHWR-LICWQIIRRGVTTIKPTVDDD-- 472  
 M. musculus -KKAARKVQKLSKNEVIVMNIIGSLSTGGRVSAVKADLGRIVITNIVCTEVG-EKIALSRRIEKHWR-LICWQIIRRGVTTIKPTVDDD-- 472  
 D. melanogaster -KKGARVEKLQKNEIIVMNIIGSLSTGGRISATKGLAKIVITTEVCTEKG-EKIALSRRIENHWR-LICWQIIPGKTIPTVLDQVA 473  
 A. thaliana -ERQMRVSKLVKGEIIVMLNIGSMSTGAMVIGVKKDMVKNITLIVCTSIG-EKVAISRIVDRHWR-LICRQIIEAGTTIFIPSPPSLA 465  
 A. fulgidus -DEEMVEKIKMNEPMLAVGATITLGVVTSARDDIVEVKKRRVVCADKG-SRVAISRIVGSRWR-LICAGIIR----- 424  
 M. thermoautotrophicum -KEETKVEPIKTGEPMLINVGTTTTIVGVVKSARADDADVVVKKLACAEEG-QRIALSRIVGARWR-LICGSIK----- 408  
 P. horikoshii -EQELRVEPIKRKEVIVLNVGTARTMGLVITNLGKDEIEVKKQIVCAEPEG-DRVAISRQIGSRWR-LICGSIKE----- 411  
 M. jannaschii -KEELKIEPLRTGEVIVMLNIGTATTAGVITSARGIADIKLKLIECAEIG-DRVAISRIVGSRWR-LICGCTIEG----- 411  
 S. solfataricus -KEMLKVDPIRAKETMLSVGSSTLGIIVTSVKDEIEVEVRRVAVVSNNI RTVISQIAGRWR-MICWQIIEI----- 415  
 P. abyssi -EQELKVEPIKRKEVIVLNVGTARTMGLVITNLGKDEIEVKKQIVCAEPEG-DRVAISRQIGSRWR-LICGSIKE----- 411

**EF1A-e/aEF1A**

T. thermophilus PQFYFRITDVTGVVQPPG-----VEMVMPGDNVFTVEIVKIVALEEG-----LRFAIREGGRTVCAVVTKILE----- 406  
 E. coli PQFYFRITDVTGTIEPEG-----VEMVMPGDNIKVVTTIHTIFAMDDG-----LRFAIREGGRTVCAVVAKVLS----- 394  
 S. cerevisiae AHIACRFDELLEKNDRRSGKGLLEDHPKFLKSGDAALVFKVPSKDMCVFAFSEYPPPLGFAVDRMQTVAVVIVKISVDKTEKAAKVTKA 451  
 S. solfataricus ASVACRVSELVSKLDPRGTQGEAKENPQFLKQGDVAIVKFKPIKELCVKEYNEFPPLGFAFMRDMGKTVGVIIIVDVKPAKVEIK---- 435  
 P. abyssi AQVPVRFEQLLAKIDPRGTNITEENPQFIKTGDSAIVVLRPMKIVVLEPVKELPQLGRFAIRDGMGMTIACMVVISIQKGE----- 428

**Annexe 1: alignement de séquence d'e/aIF2 $\gamma$  et d'EF1A-e/aIF1A réalisé avec Clustal W (Chenna *et al.*, 2003).** Les résidus parfaitement conservés entre les différentes séquences d'e/aIF2 $\gamma$  et d'EF1A-e/aEF1A sont indiqués en blanc sur fond bleu marine. Les résidus fortement conservés sont indiqués en blanc sur fond bleu clair. Les résidus parfaitement conservés pour les séquences d'e/aIF2 $\gamma$  mais non conservés pour les EF1A-e/aEF1A sont surlignés en orange. Les mutations conservatives sont indiquées en jaune. Les 4 cystéines du domaine de fixation du Zn sont indiquées en vert.

S. cerevisiae	--MSSDLAELGDFPALK-KKKKTKK-VIPDDFDAAVN-GKENGSG-----	41
S. pombe	MTETEAVVDQIDLNGALPEKKKKPKKSVAFDDEVDVAVSKDGPASSAGATDEDDSERKSSAKDLTTPVTEGEV	73
H. sapiens	-----MSGDEMI F D P T M S K K K K K K P F M L D E E G D T Q T E E T Q P S E T K E V E P E P T E D K D L E A D E E D T R K K D A S D	68
B. taurus	-----MSGDEMI F D P T M S K K K K K K P F M L D E E G D A Q T E E T Q P S E T K E V E P E P T E D K D V E A D E E D S R K K D A S D	68
M. musculus	-----MSGDEMI F D P T M S K K K K K K P F M L D E E G D A Q T E E T Q P S E T K E V E P E P T E E K D V D A D E E D S R K K D A S D	68
D. melanogaster	-----MDAEDG-FDPTLLK-KKKK-KTTFDL-----AALGLEDDTKKEDPQ-----DEASAEGGAAA	51
A. thaliana	-----MADEINEIRE-----	11
A. fulgidus	-----	
M. thermoautotrophicum	-----	
P. horikoshii	-----	
M. jannaschii	-----	
S. solfataricus	-----	
P. abyssi	-----	
S. cerevisiae	---DDLFLAGLK KKKKSKSVSADAAEAEKE-----PTDDIAEA---LGELSLKK KKKKTKDSSVDAFE	97
S. pombe	DELKDMFSSMK KKKK-SKSSASAEQE-----DITTESGE---LDFSSMK KKKKKSADLSAFE	131
H. sapiens	DLDDLNFNQK KKKKTKKIFDIDEAEEGVKDLKIESDVQEPTEPEDD---LDIMLGNK KKKKKNVKFPDEDE	139
B. taurus	DLDDLNFNQK KKKKSKKIFDIDEAEEGIKDLKIESDVQEPAEPEED---LDIMLGNK KKKKVVVKFPDEDE	139
M. musculus	DLDDLNFNQK KKKKTKKIFDIDEAEEAIAKDKIESDAQEPAEPEDD---LDIMLGNK KKKKKNVKFPDEDE	139
D. melanogaster	EEDNLDLESFG KKKKSKKPFNMDEIEAIAIPFSGGD-DVAASEPEEEEEINLDMDFSMK KKKKSKKKKELDEL	123
A. thaliana	---QEQLAPFDPS KKKKSKKVIQEPVEDLAESSQTEKSDSLPVDNGLES---SFTGMKK KKKKPTESLLNNE	79
A. fulgidus	-----	
M. thermoautotrophicum	-----	
P. horikoshii	-----	
M. jannaschii	-----	
S. solfataricus	-----	
P. abyssi	-----	
S. cerevisiae	KELAKAGLDNVDAE-SKEGTSPANS---SIQQEVGLP SELLSRF FNILRTNIP-ELAGDRSGPKFRIP PV	164
S. pombe	KELEASSTGDATSDLSKQTFDSENMGHAWLKSDRDYYP ELLNRF FTLLRTNIP-ELAGEKR--KYTIV PPS	201
H. sapiens	ILEKDEALEDEDNKKDDG-ISFSNQTPAWAGSERDYT EELLNRFVFNIMREKNPDMVAE EKR--KFVMK PPQ	208
B. taurus	VLEKDEALEDEDNKKDDG-ISFSNQTPAWAGSERDYT EELLNRFVFNIMREKNPDMVAE EKR--KFVMK PPQ	208
M. musculus	ILEKDEALEDEDNKKDDG-ISFSNQTPAWAGSERDYT EELLNRFVFNIMREKNPDMVAE EKR--KFVMK PPQ	206
D. melanogaster	FADQADDDKSEDKENDED-NSS-----TWFGSDRDYTYDELLKRVFEIILDKNPDMAAERKP--KFVMR PPQ	187
A. thaliana	SVDAGEDLDEIANDEQEGEGEIVLQQRYPWEGSERDYIYDELLGRVFNILRENNIP-ELAGDRR--RTVMR PPQ	149
A. fulgidus	-----MMR--S EEL L E R A F E K L ---A D K D V --V R R E R F E I R V S	33
M. thermoautotrophicum	-----MD--D E K L L E R A I D Q L ---P E E V F --E T K -R F E V P K A Y	31
P. horikoshii	-----M E I D Y Y D E K L L E K A Y E E L ---P E N V K --H H K S R F E V G A L	36
M. jannaschii	-----M S D L E N I D Y Y D K A L L K R A R S Q I ---P D Y V F --Q K D -R F E L E I E	39
S. solfataricus	-----M S S E K E V E M L D R L Y S K L ---P E K G R --K E G -T Q S L D N M I	34
P. abyssi	-----V R E V E I D Y Y D E K L L E K A Y E E L ---P E N V K --H H K S R F E V G A L	39
S. cerevisiae	CLRDG-KKTI FSNLQDIAEKLHRSPEHLIQ YLFAELGTS SVDGQKRLVINGKFSQKMNVLRRYILEVYTC	237
S. pombe	VHREG-KKTI FANLSDISKRMHRSLDHVIQ FLFAELGTS SVDGSSRLIIGRFQKQKQIENVLRRYIIVEVYTC	274
H. sapiens	VVRVGTKKISFVNTDICKLLHRQPKHLLA FLLAELGTS SIDGNQQLVIGRFQKQKQIENVLRRYIKEYVYTC	282
B. taurus	VVRVGTKKISFVNTDICKLLHRQPKHLLA FLLAELGTS SIDGNQQLVIGRFQKQKQIENVLRRYIKEYVYTC	282
M. musculus	VVRVGTKKISFVNTDICKLLHRQPKHLLA FLLAELGTS SIDGNQQLVIGRFQKQKQIENVLRRYIKEYVYTC	280
D. melanogaster	VLRVGTKKISFANLMDIAEALNRDPQHLLD FLLAELGTS SMDGNQQLIIGRFQKQKQIENVLRRYIKEYVYTC	261
A. thaliana	VLREGTKKISFVNTMDLCKTMHRQPDHVMQYLLAELGTS SLDGQQR LVVIGRFAPKNF GILRRYITDYVYTC	223
A. fulgidus	IQREG-ARILLKNSQIAKTLNRSSEDLHYKYIVKSLGTA GIDN-GRLVLAGKFTSESELQEVDDYVRLVYVLC	105
M. thermoautotrophicum	SVIQG-NRIFIQNFREVADALNRDPQHLLK FLLRELGTANLEG-GRALIGKFTHFLINERIEDYVNVKFIYC	103
P. horikoshii	VTIEG-NKTI IENFKDIAEALNRDPQHLLK FLLREIATATLEG-KRVVLGRFTPYLIANKLKYIKEYVYTC	108
M. jannaschii	ILIEG-NRITIRNPRELAKAVNRDEEFFAKYLLKETGSA NLEG-GRILORRISPELLKSRINDFLREVYVYTC	111
S. solfataricus	ILNIG-NTTI IIRNFAEYCDRIRREDKICMKYLLKELAAP NVDDKGE LVIGKFSQVINTLMERFLKAVYVYTC	107
P. abyssi	VTIEG-NKTI IENFKDIAEALNRDPQHLLK FLLREIATATLEG-KRVVLGRFTPYLIANK IKKYIKEYVYTC	111
S. cerevisiae	KTCKSINTELKREQSNRLFVVCXSGSTRSVSSIKTGFQAVTGKRRRM-----	285
S. pombe	KTCKSPDILLTKE--NRIFFMTEACGGSVRSVQAIKTGYQAQIGKRKHVS----	321
H. sapiens	HTCRSPDILLQKD--TRLYFLQCECHSRC SVASIKTGFQAVTGKRAQLRAKAN	333
B. taurus	HTCRSPDILLQKD--TRLYFLQCECHSRC SVASIKTGFQAVTGKRAQLRAKAN	333
M. musculus	HTCRSPDILLQKD--TRLYFLQCECHSRC SVASIKTGFQAVTGKRAQLRAKAN	331
D. melanogaster	HTCRSPDILLQKD--TRLYFLQCESCGSRC SVASIKSGFQAVTGKRAAIRAKTT	312
A. thaliana	LGCKSPDILLSKE--NRLFLLRCEKCGSQR SVAPIKTGFVARVSRKTT-----	268
A. fulgidus	RECNRPDTEFIKE--ERVMLRCLCGAKHVRNT-----	137
M. thermoautotrophicum	HECNRPDTRIIRE--GRISLLKCCACGKALKNV-----	135
P. horikoshii	PVCGSPDKIIKR--DRFYFLKCCACGAEPIQHL-----	140
M. jannaschii	RECGKPDKI I I K E --G R V H L L K C M C G A I R I R M T -----	143
S. solfataricus	STCKSLDILLKKE--KKSNIYIVCLACGAEQTVKPL-----	139
P. abyssi	PVCGSPDKI I I K R --D R F Y F L K C C A C G A E T P I Q H L -----	143

## Annexe 2: alignements de séquence d'e/aIF2 $\beta$ réalisés avec Clustal W (Chenna *et al.*, 2003).

Les résidus parfaitement conservés entre toutes les séquences d'e/aIF2 $\beta$  sont indiqués en blanc sur fond bleu. Les mutations conservatives sont indiquées en jaune. Les résidus conservés spécifiques aux protéines eucaryotiques sont surlignés en rouge et les résidus conservés spécifiques des protéines archéennes sont indiqués en vert.

S. cerevisiae	-----STSHCRFVFNKYPEIDIDVVMVNVQQLAEFCAYVKLLEYDNIEGMILLSELRRRIRSIQKLIIRVG	65
S. pombe	----MSTTSCRMENRFPEVDLVLVVMVNRQIQEFCAYVKLLEYDNIEGMVLLSELRRRIRSVQKHIRVG	66
H. sapiens	----MPGLSCRFVQHKFPEVEDEVVMVNVRSIAEFCAYVSLLEYNNIEGMILLSELRRRIRSIINKLIRIG	66
B. taurus	----PGLSCRFVQHKFPEVEDEVVMVNVRSIAEFCAYVSLLEYNNIEGMILLSELRRRIRSIINKLIRIG	65
M. musculus	----MPGLSCRFVQHKFPEVEDEVVMVNVRSIAEFCAYVSLLEYNNIEGMILLSELRRRIRSIINKLIRIG	66
D. melanogaster	----MALTSFVFNERYPEIEDEVVMVNVLSIAEFCAYVHSLLEYNNIEGMILLSELRRRIRSIINKLIRVG	65
A. thaliana	MASQTPNLECRMYEAKYPEVDMVVMVNIQVNIADFCAYVSLLEYNNIEGMILLSELRRRIRSVSSLIKVG	70
A. fulgidus	----MKEKRLIIKRSYPSKCEIVIGVVKRVLDFCAFVSLDEYFREGMVHISEVASGMIKDIRHVKKG	66
M. thermoautotrophicum	-----MFFMVRKNEWPEEGELVVGIVHKVLNFCAFATLDEYFKEAFIHISEVSSGIVKNIIRDFVREN	64
P. horikoshii	-----MPRKAREYPEEGEFVVAIVKRIHNYCAFLELDEYFKEAFMHISEVASTVVRNIRDYLKEG	61
M. jannaschii	-----MFIMRREFPEEGDIVIGIVKDVKPYCAFVELDEYFKEGMIHISEVTSGIVKNIIRHVKVG	61
S. solfataricus	-----MIYSRSLPSEGEILIAIVKQVFDYCSYVSLDEYGLQAFLPWSEVSSKIVKNIIRDVLDKEN	61
P. abyssi	-----MPRRAREYPEEGEFVVAIVKRIHNYCAFLELDEYFKEAFMHISEVASTVVRNIRDYLKEG	61
S. cerevisiae	KNDVAVVLRVMDKFKGYIDSKRRVSEDIICEEKYQKSLTVHSLIRYCAEKFQIP----LEELYKTIAW	131
S. pombe	RNECVVLRVMDKFKGYIDSKRRVSPEDVVKCEERFNKSAVHSLMRHIAEKHNVP----LETMYTTIGW	132
H. sapiens	RNECVVLRVMDKFKGYIDSKRRVPEEAIKCEDKFTKSLTVYSLRHVAEVLEYTKDEQLESLEFORTAW	135
B. taurus	RNECVVLRVMDKFKGYIDSKRRVPEEAIKCEDKFTKSLTVYSLRHVAEVLEYTKDEQLESLEFORTAW	134
M. musculus	RNECVVLRVMDKFKGYIDSKRRVPEEAIKCEDKFTKSLTVYSLRHVAEVLEYTKDEQLESLEFORTAW	135
D. melanogaster	KTEPVVLRVMDKFKGYIDSKRRVSPEDVEKTERFAKAMAINSLLRHVADILGFEGNEKLEDLYOKTAW	135
A. thaliana	RIEPVMVLRVMDKFKGYIDSKRRVSEDIQTCERYNKSILVHSLMRHVAETLSIDG----LEDLVNIGW	135
A. fulgidus	QKVICKVLDVNPKRCHIDISIKDNERQRRELQQWKNEKMAFKWLEIIEGKLNID----FKEL-EKIGK	131
M. thermoautotrophicum	QKIVARVLRVNPKRCHVDVSMKRIREDRQTKIQAWKIEQKAEKFELEAARDLGD----LDTAYEEVGY	130
P. horikoshii	QKVVAKVRVDPKRCCHIDISLRRVTVQQRKAQLQEFKRAQKAENLLRLAAEKLGKD----FEEAWREVWV	126
M. jannaschii	QRVVAKVLRVDERKCHIDISLRRVTEQQRAKQVQEWKRFQASKMLERAEEKLGS----LEEAWEVGY	126
S. solfataricus	RKVIVKVRVDRRCCTVDVSLKVVTDERRKINLQWKKIQRDLKILELVSQKLLKLS----EKDAWEQVAV	126
P. abyssi	QKVVAKVRVDPKRCCHIDISLRRVTVQQRKAQLQEFKRAQKAENLLKLAEEKLGKD----FETAWEREVWV	126
S. cerevisiae	PLSRKFGH----AYEAFKLSIID-ETVWEGIEP-----PSKDVLDLEKNIYSKRLTTP	178
S. pombe	PLYRKYGH----AYDAFKLAISNDPHVFEGLP-----PKSGVINDLLAQISRRITP	180
H. sapiens	VFDDKYKRPYGYDAFCHAVSD-PSILDSLD-----LNEDEREVLININRRLTTP	186
B. taurus	VFDDKYKRPYGYDAFCHAVSD-PSILDSLD-----LNEDEREVLININRRLTTP	185
M. musculus	VFDDKYKRPYGYDAFCHAVSD-PSILDSLD-----LNEDEREVLININRRLTTP	186
D. melanogaster	HFEKKYNN-KTVAYDIFKQSVTD-PTVDFECN-----LEPETKEVLLSNIKRKLVS	184
A. thaliana	PLYRRHGH----AFEAFKILVTDPSVGLPLTREIKEVGPDGQEVTKVVPVAVTEVVKDALVKNIARRMTP	202
A. fulgidus	KLMKEYDS----VYSAFEAAFEYEVVLAP-IVG-----EEFAKEMAIARENITP	177
M. thermoautotrophicum	ELMDFIGD----LYGAFITAAEEDEKSLIEEGVP-----EDWAAVITVAKRNITP	177
P. horikoshii	PLEEEWGE----VYAAFDAAKDGL-IEVLKGHVP-----DEWLPVLLKIVENYVEV	173
M. jannaschii	LLEDFEFG----LYNAFTMVIDEKEVLDLLEIS-----EWNKVIYEVAKESITEL	174
S. solfataricus	KLEAKYGD----PITAIKAVKESEKILIDAGVP-----EIWVKPILLKASKHAE	174
P. abyssi	PLEEEWGE----VYAAFDAAKDGL-IDVLKGHVP-----DEWLPVLLKIIDNYVEV	173
S. cerevisiae	QAVKIRADVEVSCFSYEGLDAIKDALKSAEDMSTEQMQ---VKVKLVAAELVYVLTQALDKQKGLIEQLE	243
S. pombe	QPVKIRADVEVTCFVGYEGLNAIKAAKAAEDVHTSEVP---IKVKLVAPPELVYVLTNALDKSLGLKLE	245
H. sapiens	QAVKIRADIEVACVGYEGLDAVKEALRAGLNCSTENMP---IKINLIAPERYVMTTTLERTEGLSVLS	251
B. taurus	QAVKIRADIEVACVGYEGLDAVKEALRAGLNCSTENMP---IKINLIAPERYVMTTTLERTEGLSVLN	250
M. musculus	QAVKIRADIEVACVGYEGLDAVKEALRAGLNCSTENMP---IKINLIAPERYVMTTTLERTEGLSVLN	251
D. melanogaster	PTVKIRADIEVACVGYEGLDAVKAALRAGLNCSTENMP---IRINLIAPERYVMTTSTTKKTDGLKALE	249
A. thaliana	QPKIRADIEVLCVQYDFGLVHKEAMKNAEAGNIDDCP---VKIKLVAPPELVYVLTQTLDEKQGLIILN	268
A. fulgidus	KRVKVRGYFELKSSASDGLERIKKALLEA-KKA-LVNG--VEMKLEYVGAPEKRYIVVEADYKTAENVLK	242
M. thermoautotrophicum	PEVQITGYVDIKSYAPNGVEIRKALKSAODEG-----ITVQAVGAERYLVKSTDYLYKAEKQLK	237
P. horikoshii	PTVITDAEFEITVPKPNVEIKKALIRARDRANKEKD--IEVKFTYQGAERYVVDITAPDYKAEVLE	240
M. jannaschii	TNVEVEGLVEMKSYAPDGLKQIKKALTTA-LKANPYED--VEVKITYIGAERYVVDITAPDYKAEVFK	240
S. solfataricus	RKVKMSGLITVRTNEPLGVEKIKEVISKALENIEQDYESLLNIIKYITIGAERYVVDVVGTPNKEASEALN	243
P. abyssi	PTVITDAEFEITVPKPNVEIKKALIRARDRANKEKD--VEVKFTYLGAERYVVDITAPDYKAEVLE	240
S. cerevisiae	SAIEKITEVITKYGGVCNITMPPKAVTATEDAEALQALLESKELDNRSDEDE-----	297
S. pombe	EAIKIEKISITASNGTCTVMKPKAVSTEDLELADLMKKFEKENAEISGDEE-----	299
H. sapiens	QAMAVIKEKIEEKRGVFNVMQMPKVVTTDDETELARQMERLERE-----	296
B. taurus	QAMAVIKEKIEEKRGVFNVMQMPKVVTTDDETELARQLERLERENAEVGDGDD-----	304
M. musculus	QAMAVIKEKIEEKRGVFNVMQMPKVVTTDDETELARQLERLERENAEVGDGDD-----	305
D. melanogaster	VAIEHIRAKTSEYDGEFKVIMAPKLVTIDEADLARRLERAEANAQVAGDDDEEDGADQEGMQFDPEKE	320
A. thaliana	KATAIACTETIETHKGLKVVKEGARAVSERDEKMLTEHMAKLRDLNDEEMSGDEDSGD-----	324
A. fulgidus	KATENVLKAIKKLGEGNFIREAA-----	267
M. thermoautotrophicum	EAAQKCIIEVEKEGEGEFLRELT-----	262
P. horikoshii	DIAEELRVIKQAGGEATLLRKEKIRKVKKRRK-----	275
M. jannaschii	KVCEKAVATIKKLGEGEITYRESK-----	266
S. solfataricus	QIISNLIKIGKEENVDISVVK-----	266
P. abyssi	SIAEELRVIKAEAGGEATLLRKEKIRKVKKRRK-----	275

**Annexe 3 : alignements de séquence d'e/aIF2α réalisés avec Clustal W (Chenna et al., 2003).**

Les résidus parfaitement conservés entre toutes les séquences d'e/aIF2β sont indiqués en blanc sur fond bleu. Les mutations conservatives sont indiquées en jaune. Les résidus conservés spécifiques aux protéines eucaryotiques sont surlignés en rouge et les résidus conservés spécifiques des protéines archéennes sont indiqués en vert.

# *Références Bibliographiques*



- Agrawal, R.K., Sharma, M.R., Kiel, M.C., Hirokawa, G., Booth, T.M., Spahn, C.M.T., 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. USA*, **101**, 8900-8905.
- 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.
- Anderden, G.R., Pedersen, L., Valente, L., Chatterjee, I.I., Kinzy, T.G., Kjelgaard, M. and Nyborg, J. (2000) Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A:eEF1B $\alpha$ . *Mol. Cell*, **6**, 1261-1266.
- Anderden, G.R., Nissen, P. and Nyborg, J. (2003) Elongation factors in protein biosynthesis. *Trends Biochem. Sci.*, **28**, 434-441.
- 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., Krishnamoorthy, T., Phan, L., Pavitt, G.D. and Hinnebusch, A.G. (1999) Conserved bipartite motifs in yeast eIF5 and eIF2Be, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *EMBO J.*, **18**, 1673-1688.
- 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<sup>Met</sup> is an important translation initiation intermediate *in vivo*. *Genes & Dev.*, **14**, 2534-2546.
- Asano, K., Clayton, J., Shalev, A., Phan, L., Nielsen, K., Clayton, J., Valasek, L., Donahue, T.F. and Hinnebusch, A.G. (2001) Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO J.*, **20**, 2326-2337.
- Astrom, S.U., von Pawel-Rammingen, U. and Bystrom, A.S. (1993) The yeast initiator tRNA<sup>Met</sup> can act as an elongator tRNA<sup>Met</sup> *in vivo*. *J. Mol. Biol.*, **233**, 43-58.
- Bagchi, M.K., Chakravarty, I., Datta, B., Chakrabarti, D. and Gupta, N.K. (1985) Protein synthesis in rabbit reticulocytes. A study of the mechanism of Co-eIF-2 action. *J. Biol. Chem.*, **260**, 14976-14981.
- 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.
- Barrieux, A. and Rosenfeld, M.G. (1977) Characterization of GTP-dependent Met-tRNA<sub>f</sub> binding protein. *J. Biol. Chem.*, **252**, 3843-3847.
- Barton, G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37-40.

- 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.*, **75**, 1453-1462.
- Battiste, J.L., Pestova, T.V., Hellen, C.T.U. and Wagner, G. (2000) The eIF1A solution structure reveals a large RNA-binding surface important for scanning function. *Mol. Cell*, **5**, 109-119.
- Benelli, D., Maone, E. and Londei, P. (2003) Two different mechanisms for ribosome/mRNA interaction in archaeal translation initiation. *Mol. Microbiol.*, **50**, 635-643.
- Berchtold, H., Resnikova, L., Reiser, C.O.A., Schirmer, N.K., Sprinzl, M. and Hilgenfeld, R. (1993) Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature*, **365**, 126-132.
- Biou, V., Shu, F. and Ramakrishnan, V. (1995) X-ray crystallography shows that translational initiation factor IF3 consists of two compact alpha/beta domains linked by an alpha-helix. *EMBO J.*, **14**, 4056-4064.
- Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, **7**, 1513-1523.
- Blanchard, S.C., Gonzalez Jr., R.L., Kim, H.D., Chu, S. and Puglisi, J.D. (2004) tRNA selection and kinetic proofreading in translation. *Nat. Struct. Mol. Biol.*, **11**, 1008-1014.
- Boelens, R. and Gualerzi, C.O. (2002) Structure and Function of Bacterial Initiation Factors. *Current Protein and Peptide Science*, **3**, 107-119.
- Bommer, U.A., Salimans, M.M., Kurzchalia, T.V., Voorma, H.O. and Karpova, G.G. (1988) Affinity labeling by a photoreactive GTP analogue of the alpha-subunit of eukaryotic initiation factor eIF-2 in different initiation complexes. *Biochem. Int.*, **16**, 549-557.
- Bommer, U.A. and Kurzchalia, T.V. (1989) GTP interacts through its ribose and phosphate moieties with different subunits of the eukaryotic initiation factor eIF-2. *FEBS Letters*, **244**, 323-327.
- Bommer, U.A., Kraft, R., Kurzchalia, T.V., Price, N.T. and Proud, C.G. (1991) Amino acids sequence analysis of the beta- and gamma-subunits of eukaryotic initiation factor eIF-2. Identification of regions interacting with GTP. *Biochim. Biophys. Acta*, **1079**, 308-315.
- 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.
- Bourne, H.R., Sanders, D.A. and McCormick, F. (1991) The GTPase superfamily : conserved structure and molecular mechanism. *Nature*, **349**, 117-127.

- Brevet, A., Chen, J., Lévêque, F., Plateau, P. and Blanquet, S. (1989) In vivo synthesis of adenylylated bis(5'-nucleosidyl)tetraphosphates (Ap<sub>4</sub>N) by *Escherichia coli* aminoacyl-tRNA synthetases. *Proc. Natl. Acad. Sci. USA*, **86**, 8275-8279.
- Brock, S., Szkaradkiewicz, K. and Sprinzl, M. (1998) Initiation factors of protein biosynthesis in bacteria and their structural relationship to elongation and termination factors. *Mol. Microbiol.*, **29**, 409-417.
- Brunger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta. Crystallogr. D54*, 905-921.
- Canonaco, M.A., Calogero, R.A. and Gualerzi, C.O. (1986) Mechanism of translational initiation in prokaryotes. Evidence for a direct effect of IF2 on the activity of the 30S ribosomal subunit. *FEBS Letters*, **207**, 198-204.
- Carter, A.P., Clemons Jr., W.M., 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., Jonhson, 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, J. 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.
- Castilho-Valavicius, B., Thompson, G.M. and Donahue, T.F. (1992) Mutation analysis of the Cys-X<sub>2</sub>-Cys-X<sub>19</sub>-Cys-X<sub>2</sub>-Cys motif on the β subunit of eukaryotic translation initiation factor 2. *Gene Expr.*, **2**, 297-309.
- 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., Offenhäuser, N., Marchisio, P.C. and Biffo, S. (2003) Release of eIF6 (p27<sup>BBP</sup>) from the 60S subunit allows 80S ribosome assembly. *Nature*, **426**, 579-584.
- Chaudhuri, J., Chowdhury, D. And Maitra, U. (1999) Distinct functions of eukaryotic initiation factors eIF1A and eIF3 in the formation of the 40S ribosomal preinitiation complex. *J. Biol. Chem.*, **274**, 17975-17980.
- Chenna, R., Sugarawa, 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.
- Cherfils, J. and Chardin, P. (1999) GEFs: structural basis for their activation of small GTP-binding proteins. *Trends Biochem. Sci.*, **24**, 306-311.

- 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 eIF2 $\beta$ /eIF5 family of proteins. *Biochemistry*, **41**, 5730-5742.
- Cigan, A.M., Pabich, E.K., Feng, L. and Donahue, T.F. (1989) Yeast translation initiation suppressor *sui2* encodes the  $\alpha$  subunit of eukaryotic initiation factor 2 and shares sequence identity with the human  $\alpha$  subunit. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 2784-2788.
- Clemons Jr., W.M., Brodersen, D.E., McCutcheon, J.P., May, J.L., Carter, A.P., Morgan-Warren, R.J., Wimberly, B.T. and Ramakrishnan, V. (2001) Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: purification, crystallization and structure determination. *J. Mol. Biol.*, **310**, 827-843.
- Collaborative computational project No.4 (1994) The CCP4 suite: programs for protein crystallography. *Acta. Crystallogr. D50*, 760-763.
- Colthurst, D.R. and Proud, C.G. (1986) Eukaryotic initiation factor 2 from rat liver : no apparent function for the beta-subunit in the formation of initiation complexes. *Biochim. Biophys. Acta*, **868**, 77-86.
- Condo, I., Ciammaruconi, A., Benelli, D., Ruggero, D. and Londei, P. (1999) Cis-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.*, **34**, 377-384.
- Cousineau, B., Leclerc, F. and Cedergren, R. (1997) On the origin of protein synthesis factors: A gene duplication/fusion model. *J. Mol. Evol.*, **45**, 661-670.
- Cowtan, K.D. and Main, P. (1993) Improvement of macromolecular electron density maps by the simultaneous application of real and reciprocal space constraints. *Acta. Crystallogr. D49*, 148-157.
- Culver, G.M. (2001) Meanderings of the mRNA through the ribosome. *Structure*, **9**, 751-758.
- Dahlquist, K.D. And Puglisi, J.D. (2001) Interaction of 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.
- 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.
- Dardel, F. (1994) MC-Fit : Using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Comput. Appl. Biosci.*, **10**, 273-275.
- 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.
- 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.
- 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.
- Dever, T.E. (2002) Gene-specific regulation by general translation factors. *Cell*, **108**, 545-556.
- Dey, D., Oleinikov, A.V. and Traut, R.R. (1995) The hinge region of *Escherichia coli* ribosomal protein L7/L12 is required for factor binding and GTP hydrolysis. *Biochimie*, **77**, 925-930.
- Dhaliwal, S. and Hoffman, D.W. (2003) The crystal structure of the N-terminal region of the alpha subunit of translation initiation factor 2 (eIF2alpha) from *Saccharomyces cerevisiae* provides a view of the loop containing serine 51, the target of the eIF2alpha-specific kinases. *J. Mol. Biol.*, **334**, 187-195.
- Donahue, T.F., Cigan, A.M., Pabich, E.K. and Valavicius, B.C. (1988) Mutations at a Zn(II) finger motif in the yeast eIF-2 $\beta$  gene alter ribosomal start-site selection during the scanning process. *Cell*, **54**, 621-632.
- Dorris, D.R., Erickson, F.L. and Hannig, E.M. (1995) Mutations in *GCD11*, the structural gene for eIF2 $\gamma$  in yeast, alter translational regulation of *GCN4* and the selection of the start site for protein synthesis. *EMBO J.*, **10**, 2239-2249.
- Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E.coli* by high voltage electroporation. *Nucleic Acids Res.*, **16**, 6127-6145.
- Erickson, F.L. and Hannig, E.M. (1996) Ligand interactions with eukaryotic translation initiation factor 2: role of the  $\gamma$ -subunit. *EMBO J.*, **15**, 6311-6320.
- Erickson, F.L., Harding, L.D., Dorris, D.R. and Hannig, E.M. (1997) Functional analysis of homologs of translation initiation factor 2  $\gamma$  in yeast. *Mol. Gen. Genet.*, **253**, 711-719.
- Evans, S.V. (1993) Setor: hardware lighted three-dimensional solid model representation of macromolecules. *J. Mol. Graphics*, **11**, 134-138.
- Farruggio, D., Chaudhuri, J., Maitra, U. and RajBhandary, U.L. (1996) The A1  $\times$  U72 base pair conserved in eukaryotic initiator tRNAs is important specifically for binding to eukaryotic translation initiation factor eIF2. *Mol. Cell. Biol.*, **16**, 4248-4256.
- Fletcher, C.M., Pestova, T.V., Hellen, C.T.U. and Wagner, G. (1999) Structure and interactions of the translation initiation factor eIF1. *EMBO J.*, **18**, 2631-2637.

- Flynn, A., Oldfield, S. and Proud, C.G. (1993) The role of the beta-subunit of initiation factor eIF-2 in initiation complex formation. *Biochim. Biophys. Acta*, **1174**, 117-121.
- Flynn, A., Shatsky, I.N., Proud, C.G. and Kaminski, A. (1994) The RNA-binding properties of protein synthesis initiation factor eIF-2. *Biochim. Biophys. Acta*, **1219**, 293-301.
- Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., Lata, R.K. and Agrawal, R.K. (1995) A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome. *Nature*, **376**, 441-444.
- Fraser, C.S., Lee, J.Y., Mayeur, G.L., Bushell, M., Doudna, J.A. and Hershey, J.W.B. (2004) The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40S ribosomal subunits *in vitro*. *J. Biol. Chem.*, **279**, 8946-8956.
- Futterer, J., Kiss-Laszlo, Z. and Hohn, T. (1993) Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell*, **73**, 789-802.
- 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.
- Gaspar, N.J., Kinzy, T.G., Scherer, B.J., Humbelin, M., Hershey, J.W.B. and Merrick, W.C. (1994) Translation initiation factor eIF-2. Cloning and expression of the human cDNA encoding the gamma-subunit. *J. Biol. Chem.*, **269**, 3415-3422.
- Gonsky, R., Itamar, D., Harary, R. and Kaempfer, R. (1992) Binding of ATP and messenger RNA by the  $\beta$ -subunit of eukaryotic initiation factor 2. *Biochimie*, **74**, 427-434.
- Gribskov, M. (1992) Translation initiation factor IF1 and factor eIF2 $\alpha$  share an RNA-binding motif with prokaryotic ribosomal S1 and polynucleotide phosphorylase. *Gene*, **119**, 107-111.
- Grigoriadou, C., Marzi, S., Seo, H.S., Dongli, P., Kirillov, S., Gualerzi, C.O. and Cooperman, B.S. (2004) In *Translational Control* (A. Hinnebusch, T. Pestova, and J. Richter, Eds.) IF2, IF3 and thiostrepton interaction during initiation complex formation., pp. 407, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).
- Groft, C.M., Beckmann, R., Sali, A. and Burley, S.K. (2000) Crystal structures of ribosome anti-association factor IF6. *Nat. Struct. Biol.*, **7**, 1156-1163.
- 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.O., 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.
- Gualerzi, C.O. and Pon, C.L. (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry*, **29**, 5881-5889.
- Gualerzi, C.O., Brandi, L., Caserta, E., La Teana, A., Spurio, R., Tomsic, J. and Pon, C.L. (2000) in *The Ribosome: structure, function, antibiotics, and cellular interactions* (Garret, R.A., Douthwaite, S. R., Liljas, A., Matheson, A. T., Moore, P. B. and Noller, H. F.) Translation initiation in bacteria, pp. 477-494, ASM Press, Washington, D.C.
- Guenneugues, M., Caserta, E., Brandi, L., Spurio, R., Meunier, S., Pon, C.L., Boelens, R. and Gualerzi, C.O. (2000) Mapping the fMet-tRNA<sup>Met</sup> binding site of initiation factor IF2. *EMBO J.*, **19**, 2533-5240.
- Guillon, J.M., Meinnel, T., Mechulam, Y., Lazennec, C., Blanquet, S. and Fayat, G. (1992b) Nucleotides of tRNA governing the specificity of *Escherichia coli* methionyl-tRNA<sup>Met</sup> formyltransferase. *J. Mol. Biol.*, **224**, 359-367.
- 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., 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, L., Schmitt, E., Blanquet, S. and Mechulam, Y. (2005) Initiator tRNA binding e/aIF5B, the eukaryotic/archaeal homologue of bacterial initiation factor IF2. *Biochemistry*, **44**, 15594-15601.
- Gutiérrez, 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 aIF2 $\beta$ . *FEBS Letters*, **517**, 155-158.
- Gutiérrez, P., Osborne, M.J., Siddiqui, N., Trempe, J.F., Arrowsmith, C. and Gehring, K. (2004) Structure of the archaeal translation initiation factor aIF2 $\beta$  from *Methanobacterium thermoautotrophicum* : implications for translation initiation. *Protein Science*, **13**, 659-667.
- 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.
- Hartz, D., McPheeters, D.S. and Gold, L. (1989) Selection of the initiator tRNA by *Escherichia coli* initiation factors. *Genes & Dev.*, **3**, 1899-1912.

- 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.
- Hashimoto, N.N., Carnevalli, L.S. and Castilho, B.A. (2002) Translation initiation at non-AUG codons mediated by weakened association of eIF2 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.
- Hershey, J.W.B. and Merrick, W.C. (2000) in *Translational control of gene expression* (Sonenberg, N., Hershey, J. W. B. and Mathews, M. B., Eds.) The pathway and mechanism of initiation of protein synthesis, pp.33-88, Cold Spring Harbor Laboratory Press, Plainview, NY
- Hinnebusch, A.G. (2000) in *Translational control of gene expression* (Sonenberg, N., Hershey, J. W. B. and Mathews, M. B., Eds.) Mechanism and regulation of initiator methionyl-tRNA binding to ribosomes, pp. 185-244, Cold Spring Harbor Laboratory Press, Plainview, NY.
- Hirel, P.H., Lévêque, F., Mellot, P., Dardel, F., Panvert, M., Mechulam, Y. and Fayat, G. (1988) Genetic engineering of methionyl-tRNA synthetase: *in vitro* regeneration of an active synthetase by proteolytic cleavage of a methionyl-tRNA synthetase- $\beta$ -galactosidase chimeric protein. *Biochimie*, **70**, 773-782.
- 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.
- Iborra, F.J., Escargueil, A.E., Kwek, K.Y., Akoulitchev, A. and Cook, P.R. (2004) Molecular cross-talk between the transcription, translation, and nonsense-mediated decay machineries. *J. Cell Sci.*, **117**, 899-906.
- Iost, I., Dreyfus, M. and Linder, P. (1999) Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. *J. Biol. Chem.*, **274**, 17677-17683.
- Ito, T., Marintchev, A. and Wagner, G. (2004) Solution structure of human initiation factor eIF2 $\alpha$  reveals homology to the elongation factor eEF1B. *Structure (Camb)* **12**, 1693-1704.
- Jeppesen, M.G., Navratil, T., Spremulli, L.L. and Nyborg, J. (2005) Crystal structure of the bovine mitochondrial elongation factor Tu.Ts complex. *J. Biol. Chem.*, **280**, 5071-5081.
- Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for the building of proteins model in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110-119.

- Jue, R.A., Woodbury, N.W. and Doolittle, R.F. (1980) Sequence homologies among E. Coli ribosomal proteins: Evidence for evolutionarily related groupings and internal duplications. *J. Mol. Evol.*, **15**, 129-148.
- Kabsch, W. (1988) Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Cryst.*, **21**, 916-924.
- Kahvejian, A., Svitkin, Y.V., Sukarieh, R., M'Boutchou, M.N. and Sonenberg, N. (2005) Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes & Dev.*, **19**, 104-113.
- Kainuma, M. and Hershey, J.W.B. (2001) Depletion and deletion analysis of eukaryotic translation initiation factor 1A in *Saccharomyces cerevisiae*. *Biochimie*, **83**, 505-514.
- Kang, H.A. and Hershey, J.W. (1994) Effect of initiation factor eIF-5A depletion on protein synthesis and proliferation of *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **269**, 3934-3940.
- Kapp, L.D. and Lorsch, J.R. (2003) 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. (2004) The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem.*, **73**, 657-704.
- Karimi, R., Pavlov, M.Y., Buckingham, R. and Ehrenberg, M. (1999) Novel roles for classical factors at the interface between translation termination and initiation. *Mol. Cell*, **3**, 601-609.
- Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. and Leberman, R. (1996) The structure of the *Escherichia coli* EF-Tu.EF-Ts complex at 2.5 Å resolution. *Nature*, **379**, 511-518.
- Keeling, P.J. Keeling, P.J. and Doolittle, W.F. (1995) Archaea: Narrowing the gap between prokaryotes and eukaryotes. *Proc. Natl. Acad. Sci. USA*, **92**, 5761-5764.
- Keeling, P.J., Fast, N. and McFadden, G.I. (1998) Evolutionary relationship between translation initiation factor eIF-2 $\gamma$  and selenocysteine-specific elongation factor SelB: change of function in translation factors. *J. Mol. Evol.*, **47**, 649-655.
- Kennell, D. and Riezman, H. (1977) Transcription and translation initiation frequencies of the *Escherichia coli lac* operon. *J. Mol. Biol.*, **304**, 219-229.
- Kimball, S.R., Heinzinger, N.K., Horetsky, R.L. and Jefferson, L.S. (1998) Identification of interprotein interactions between the subunits of eukaryotic initiation factors eIF2 and eIF2B. *J. Biol. Chem.*, **273**, 3039-3044.
- Kimball, S.R., Fabian, J.R., Pavitt, G.D., Hinnebusch, A.G. and Jefferson, L.S. (1998) Regulation of guanine nucleotide exchange through phosphorylation of eukaryotic initiation factor eIF2 $\alpha$ . *J. Biol. Chem.*, **273**, 12841-12845.

- Kimball, S.R. (1999) Eukaryotic initiation factor eIF2. *Int. J. Biochem. Cell Biol.*, **31**, 25-29.
- Kjeldgaard, M., Nissen, P., Thirup, S. and Nyborg, J. (1993) The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure*, **1**, 35-50.
- Kolupaeva, V.G., Unbehauen, A., Lomakin, I.B., Hellen, C.T.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.
- Koonin, E.V. (1995) Multidomain organization of eukaryotic guanine nucleotide exchange translation initiation factor eIF-2B subunits revealed by analysis of conserved sequence motifs. *Protein Science*, **4**, 1608–1617.
- Korneeva, N.L., Lamphear, B.J., Hennigan, F.L. and Rhoads, R.E. (2000) Mutually cooperative binding of eukaryotic translation initiation factor eIF3 and eIF4A to human eIF4G-1. *J. Biol. Chem.*, **275**, 41369-41376.
- Kozak, M. (1994) Determinants of translational fidelity and efficiency in vertebrate mRNAs. *Biochimie*, **76**, 815-821.
- Kozak, M. (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene*, **299**, 1-34.
- 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 (eIF2 $\alpha$ ) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. *Mol. Cell. Biol.*, **21**, 5018-5030.
- Kyripides, N.C. and Woese, C.R. (1998) Universally conserved translation initiation factors. *Proc. Natl. Acad. Sci. USA*, **95**, 224–228.
- Kyripides, N.C. and Woese, C.R. (1998) Archaeal translation initiation revisited : the initiation factor 2 and eukaryotic initiation factor 2B alpha-beta-delta subunit families. *Proc. Natl. Acad. Sci. USA*, **95**, 3726-3730.
- Laalami, S., Putzer, H., Plumbridge, J.A. and Grunberg-Manago, M. (1991) A severely truncated form of translational initiation factor 2 supports growth of *Escherichia coli*. *J. Mol. Biol.*, **220**, 335-349.
- La Cour, T.F., Nyborg, J., Thirup, S. and Clark, B.F. (1985) Structural details of the binding of guanosine diphosphate to elongation factor Tu from *E. coli* as studied by X-Ray crystallography. *EMBO J.*, **4**, 2385-2388.
- Lakowski, R.A., Mac Arthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Crystallogr.*, **26**, 283-291.
- Lamphear, B.J., Kirchweger, R., E., 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.*, **268**, 19200-19203.
- 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.
- Laurino, J.P., Thompson, G.M., Pacheco, E. and Castilho, B.A. (1999) The  $\beta$  subunit of eukaryotic translation initiation factor 2 binds mRNA through the lysine repeats and a region comprising the C<sub>2</sub>-C<sub>2</sub> motif. *Mol. Cell. Biol.*, **19**, 173-181.
- Laursen, B.S., Mortensen, K.K., Sperling-Petersen, H.U. and Hoffman, D.W. (2003) A conserved structural motif at the N terminus of bacterial translation initiation factor IF2. *J. Biol. Chem.*, **278**, 16320-16328.
- Laursen, B.S., Sørensen, H.P., Mortensen, K.K. and Sperling-Petersen, H.U. (2005) Initiation of Protein Synthesis in Bacteria. *Microbiol. Mol. Biol. Rev.*, **69**, 101-123.
- Lecompte, O., Ripp, R., Puzos-Barbe, V., Duprat, S., Heilig, R., Dietrich, J., Thierry J-C. and Poch O. (2001) Genome evolution at the genus level: comparison of three complete genomes of hyperthermophilic archaea.
- 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. USA*, **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. USA*, **99**, 16689-16694.
- Leibundgut, M., Frick, C., Thanbichler, M., Böck, A. and Ban, N. (2005) Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *EMBO J.*, **24**, 11-22.
- Leslie, A. (1990) in *Crystallographic computing V*, Oxford University Press.
- 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 Science*, **10**, 2426-2438.

- Lockwood, A.H., Sarkar, P. and Maitra, U. (1972) Release of polypeptide chain initiation factor IF-2 during initiation complex formation. *Proc. Natl. Acad. Sci. USA*, **69**, 3602–3605.
- 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.
- Londei, P. (2005) Evolution of translational initiation : new insights from the archaea. *FEMS Microbiol. Rev.*, **29**, 185-200.
- Lorsch, J.R. and Herschlag, D. (1998) The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochemistry*, **37**, 2194-206.
- Lorsch, J.R. and Herschlag, D. (1999) Kinetic dissection of fundamental processes of eukaryotic translation in vitro. *EMBO J.*, **18**, 6705-6717.
- Luchin, S., Putzer, H., Hershey, J.W., Cenatiempo, Y., Grunberg-Manago, M. and Laalami, S. (1999) In vitro study of two dominant inhibitory GTPase mutants of *Escherichia coli* translation initiation factor IF2. Direct evidence that GTP hydrolysis is necessary for factor recycling. *J. Biol. Chem.*, **274**, 6074-6079.
- Maag, D., Fekete, C.A., Gryczynski, A. and Lorsch, J.R. (2005) A conformational change in the eucaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. *Mol. Cell*, **17**, 265-275.
- Majumdar, A., Bose, K.K. and Gupta, N.K. (1976) Specific binding of *Escherichia coli* chain initiation Factor 2 to fMet-tRNA<sup>Met</sup>. *J. Biol. Chem.*, **251**, 137-140.
- Mangroo, D., Wu, X.Q. and RajBhandary, U.L. (1995) *Escherichia coli* initiator tRNA: structure-function relationships and interactions with the translational machinery. *Biochem. Cell Biol.*, **73**, 1023–1031.
- 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 (N.Y.)*, **8**, 1189-1232.
- Marcotrigiano, J., Gingras, A.C., Soneneberg, 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.
- 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. USA*, **100**, 1535-1540.

- Marzi, S., Knight, W., Brandi, L., Caserta, E., Soboleva, N., Hill, W.E., Gualerzi, C.O. and Lodmell, J.S. (2003) Ribosomal localization of translation initiation factor IF2. *RNA*, **9**, 958–969.
- Mayaux, J-F., Kalogerakos, T., Brito, K.K. and Blanquet, S. (1982) Removal of the tightly bound zinc from *Escherichia coli* trypsin-modified methionyl-tRNA synthetase. *Eur. J. Biochem.*, **128**, 41-46.
- Mayer, C., Köhrer, C., Kenny, E., Prusko, C. and RajBhandary, U.T. (2003) Anticodon sequence mutants of *Escherichia coli* initiator tRNA : effects of overproduction of aminoacyl-tRNA synthetases, methionyl-tRNA formyltransferase, and initiation factor 2 on activity in initiation. *Biochemistry*, **42**, 4787-4799.
- Mazza, C., Segref, A., Mattaj, I.W. and Cusack, S. (2002) Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.*, **21**, 5548-5557.
- McCutcheon, J.P., Agrawal, R.K., Philips, S.M., Grassucci, R.A., Gerchman, S.E., Clemons Jr., W.M., Ramakrishnan, V. and Frank, J. (1999) Location of translational initiation factor IF3 on the small ribosomal subunit. *Proc. Natl. Acad. Sci. USA*, **96**, 4301–4306.
- 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. *J. Appl. Cryst.*, **35**, 538-545.
- Medhi, Q. and Yudkin, M.D. (1968) Coupling of transcription to translation in the synthesis of enzymes by *Escherichia coli*. *Biochem. J.*, **109**, 2936.
- Meinzel, T., Mechulam, Y. and Fayat, G. (1988) Fast purification of a functional elongator tRNA<sup>Met</sup> expressed from a synthetic gene in vivo. *Nucleic Acids Res.*, **16**, 8095-8096.
- Meinzel, T., Mechulam, Y., Lazennec, C., Blanquet, S. and Fayat, G. (1993) Critical role of the acceptor stem of tRNAs(Met) in their aminoacylation by *Escherichia coli* methionyl-tRNA synthetase. *J. Mol. Biol.*, **229**, 26-36.
- Meinzel, T. and Blanquet, S. (1995) Maturation of pre-tRNA<sup>Met</sup> by *E. coli* Rnase P is specified by a guanosine of the 5' flanking sequence. *J. Biol. Chem.*, **270**, 15906-15914.
- 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**, 429-433.
- Meunier, S., Spurio, R., Czisch, M., Wechselberger, R., Guenneugues, M., Gualerzi, C.O. and Boelens, R. (2000) Structure of the to fMet-tRNA<sup>Met</sup>-binding domain of *B. stearothermophilus* initiation factor IF2. *EMBO J.*, **19**, 1918-1926.

- Miller, J.F. (1992) *A short course in bacterial genetics*. Cold Spring Harbor Laboratory Press, New York.
- Mitsui, K-I., Datta, A. and Ochoa, S. (1981) Removal of the  $\beta$  subunit of the eukaryotic polypeptide chain initiation factor 2 by limited proteolysis. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 4128-4132.
- Mohr, D., Wintermeyer, W. and Rodnina, M.V. (2002) GTPase activation of elongation factors Tu and G on the ribosome. *Biochemistry*, **41**, 12520–12528.
- Moreno, J.M.P., Kildsgaard, J., Siwanowicz, I., Mortensen, K.K. and Sperling-Petersen, H.U. (1998) Binding of *Escherichia coli* initiation factor IF2 to 30S ribosomal subunits: a functional role for the N-terminus of the factor. *Biochem. Biophys. Res. Commun.*, **252**, 465-471.
- Moreno, J.M.P., Drskjotersen, L., Kristensen, J.E., Mortensen, K.K. and Sperling-Petersen, H.U. (1999) Characterization of the domains of *E. coli* initiation factor IF2 responsible for recognition of the ribosome. *FEBS Letters*, **455**, 130–134.
- Morley, S.J., Curtis, P.S. and Pain, V.M. (1997) eIF4G: translation's mystery factor begins to yield its secrets. *RNA*, **3**, 1085-1104.
- Mortensen, K.K., Kildsgaard, J., Moreno, J.M., Steffensen, S.A., Egebjerg, J. and Sperling-Petersen, H.U. (1998) A six-domain structural model for *Escherichia coli* translation initiation factor IF2. Characterisation of twelve surface epitopes. *Biochem. Mol. Biol. Int.*, **46**, 1027–1041.
- Mosyak, L., Reshetnikova, L., Goldur, Y., Delarue, M. and Safro, M.G. (1995) Structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus*. *Nat. Struct. Biol.*, **2**, 537-547.
- Murzin, A.G. (1993) OB(oligonucleotide / oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861-867.
- Nagai, K., Oubridge, C., Jessen, T.H., Li, J. and Evans, P.R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515-520.
- Naranda, T., Sirangelo, I., Fabbri, B.J. and Hershey, J.W.B. (1995) Mutations in the NKXD consensus element indicate that GTP binds to the  $\gamma$ -subunit of translation initiation factor eIF2. *FEBS Letters*, **372**, 249-252.
- Nicholls, A. and Honig, B. (1991) A rapid finite difference algorithm utilizing successive over-relaxation to solve the Poisson-Boltzmann equation, *J. Comput. Chem.*, **12**, 435-445.
- Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins*, **11**, 281-286.

- Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszynska, A., Dadlez, M., Gingras, A.C., Mak, P., Darzynkiewicz, E., Sonenberg, N., Burley, S.K. and Stolarski, R. (2002) Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J. Mol. Biol.*, **319**, 615-635.
- Nika, J., Rippel, S. and Hannig, E.M. (2001) Biochemical analysis of the eIF2 $\beta\gamma$  complex reveals structural function for eIF2 $\alpha$  in catalyzed nucleotide exchange. *J. Biol. Chem.*, **276**, 1051-1056.
- Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B.F.C. and Nyborg, J. (1995) Crystal structure of the ternary complex of Phe-tRNA<sup>Phe</sup>, EF-Tu, and a GTP analog. *Science*, **270**, 1464-1472.
- Nissen, P., Thirup, S., Kjeldgaard, M. and Nyborg, J. (1999) The crystal structure of Cys-tRNA<sup>Cys</sup>-EF-Tu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Structure*, **7**, 143-156.
- 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.
- Nonato, M.C., Widom, J. and Clardy, J. (2002) Crystal structure of the N-terminal segment of human eukaryotic translation initiation factor 2 alpha. *J. Biol. Chem.*, **277**, 17057-17061.
- Nygaard, O., Westermann, P. and Hultin, T. (1980) Met-tRNA-Met-f is located in close proximity to the beta subunit of eIF-2 in the eukaryotic initiation complex, eIF-2.Met-tRNA-Met-f.GDP-PCP. *FEBS Letters*, **113**, 125-128.
- Ogle, J.M., Carter, A.P. and Ramakrishnan, V. (2003) Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem. Sci.*, **28**, 259-266.
- 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.
- Olsthoorn, R.C., Zoog, S. and van Duin, J. (1995) Coevolution of RNA helix stability and Shine-Dalgarno complementarity in a translational start region. *Mol. Microbiol.*, **15**, 333-339.
- Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W. and Wittinghofer, A. (1990) Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution : implications for the mechanism of GTP hydrolysis. *EMBO J.*, **9**, 2351-2359.
- Pain, V.M. (1996) Initiation of protein synthesis in eukaryotic cells. *Eur. J. Biochem.*, **236**, 747-771.
- Pavitt, G.D., Yang, W. and Hinnebusch, A.G. (1997) Homologous segments in three subunits of the guanine nucleotide exchange factor eIF2B mediate translational regulation by phosphorylation of eIF2. *Mol. Cell. Biol.*, **17**, 1298-1313.

- Pavitt, G.D., Ramaiah, K.V.A., 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.
- Pedullà, N., Palermo, R., Hasenöhrl, D., Bläsi, U., Cammarano, P. and Londei, P. (2005) The archaeal eIF2 homologue: functional properties of an ancient translation initiation factor. *Nucleic Acids Res.*, **33**, 1804-1812.
- 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., 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. 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.
- 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<sub>f</sub><sup>Met</sup>. *Biochem. Biophys. Res Commun.*, **91**, 1068-1074.
- Petrelli, D., La Teana, A., Garofalo, C., Spurio, R. Pon, C.L. and Gualerzi, C.O. (2001) Translation initiation factor IF3 : two domains, five functions, one mechanism ? *EMBO J.*, **20**, 4560-4569.
- 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<sub>i</sub><sup>Met</sup>. *EMBO J.*, **20**, 2954-2965.
- Picciano, D.J., Prichard, P.M., Merrick, W.C., Shafritz, D.A., Graf, H., Crystal, R.G. and Anderson, W.F. (1973) Isolation of protein synthesis initiation factors from rabbit liver. *J. Biol. Chem.*, **248**, 204-214.
- Pingoud, A., Urbanke, C., Krauss, G., Peters, F. and Maass, G. (1977) Ternary complex formation between elongation factor Tu, GTP and aminoacyl-tRNA: an equilibrium study. *Eur. J. Biochem.*, **78**, 403-409.
- Pioletti, M., Schlunzen, F., Harms, J., Zarivach, R., Gluhmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A. and Franceschi, F. (2001) Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *EMBO J.*, **20**, 1829-1839.
- Pisarev, A.V., Shirokikh, N.E. and Hellen, C.U. (2005) Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites. *C. R. Biol.*, **328**, 589-605.

- Polekhina, G., Thirup, S., Kjølgaard, M., Nissen, P., Lippmann, C. and Nyborg, J. (1996) Helix unwinding in the effector region of elongation factor EF-Tu-GDP. *Structure*, **4**, 1141-1151.
- Pon, C.L. and Gualerzi, C.O. (1984) Mechanism of protein biosynthesis in procaryotic cells. Effect of initiation factor IF1 on the initial rate of 30S initiation complex formation. *FEBS Letters*, **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.
- Preisse, T. and Hentze, M.W. (2003) Starting the protein synthesis machine: eukaryotic translation initiation. *BioEssays*, **25**, 1201-1211.
- Proweller, A. and Butler, J.S. (1996) Ribosomal association of poly(A)-binding protein in poly(A)-deficient *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **271**, 10859-10865.
- Ramakrishnan, V. (2002) Ribosome structure and the mechanism of translation. *Cell*, **108**, 557-572.
- Ramesh, V. and RajBhandary, U.L. (2001) Importance of the anticodon sequence in the aminoacylation of tRNAs by methionyl-tRNA synthetase and by valyl-tRNA synthetase in an Archaeobacterium. *J. Biol. Chem.*, **276**, 3660-3665.
- Rogers Jr, G.W., Richter, N.J. and Merrick, W.C. (1999) Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J. Biol. Chem.*, **274**, 12236-12244.
- Rogers Jr, G.W., 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.
- Roll-Mecak, A., Alone, P., Cao, C., Dever, T.E. and Burley, S.K. (2004) X-ray structure of translation initiation factor eIF2 $\gamma$ : implications for tRNA and eIF2 $\alpha$  binding. *J. Biol. Chem.*, **279**, 10634-10642.
- Rosenzweig, A.C., Huffman, D.L., Hou, M.Y., Wernimont, A.K., Pufahl, R.A. and O'Halloran, T.V. (1999) Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. *Structure*, **7**, 605-617.
- Rudinger, J., Hillenbrandt, R., Sprinzl, M. and Gieger, R. (1996) Antideterminants present in minihelix(Sec) hinder its recognition by prokaryotic elongation factor Tu. *EMBO J.*, **15**, 650-657.

- Samuel, C.E. (1993) The eIF-2  $\alpha$  protein kinases, regulators of translation in eukaryotes from yeasts to humans. *J. Biol. Chem.*, **268**, 7603-7606.
- 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.
- Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmüller, L., Lautwein, A., Schmitz, F. and Wittinghofer, A. (1997) The Ras-RasGAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, **277**, 333-338.
- Scheffzek, K., Ahmadian, M.R. and Wittinghofer, A. (1998) GTPase-activating proteins: helping hands to complement an active site. *Trends Biochem. Sci.*, **23**, 257-262.
- Schenk, P.M., Baumann, S., Mattes, R. and Steinbiss, H.H. (1995) Improved high-level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare Arg tRNAs. *Biotechniques*, **19**, 196-200.
- Schlutzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (2000) Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell*, **102**, 615-623.
- Schmitt, E., Panvert, M., Blanquet, S. and Mechulam, Y. (1998) Crystal structure of methionyl-tRNA<sup>Met</sup> transformylase complexed with the initiator formyl-methionyl-tRNA<sup>Met</sup>. *EMBO J.*, **17**, 6819-6826.
- Schmitt, E., Blanquet, S. and Mechulam, Y. (2002) The large subunit of initiation factor aIF2 is a close structural homolog of elongation factors. *EMBO J.*, **21**, 1821-1832.
- Schneider, T.D., Stormo, G.D., Gold, L. and Ehrenfeucht, A. (1986) Information content of binding sites on nucleotide sequences. *J. Mol. Biol.*, **188**, 415-431.
- Schulman, L.H. and Pelka, H. (1988) Anticodon switching changes the identity of methionine and valine transfer RNAs. *Science*, **242**, 765-768.
- Sengupta, J., Agrawal, R.K. and Frank, J. (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc. Natl. Acad. Sci. USA*, **98**, 11991-11996.
- 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.
- She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., Chan-Weiher, C.C., Clausen, I.G., Curtis, B.A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P.M., Heikamp-de Jong, I., Jeffries, A.C., Kozera, C.J., Medina, N., Peng, X., Thi-Ngoc, H.P., Redder, P., Schenk, M.E., Theriault, C., Tolstrup, N., Charlebois, R.L., Doolittle, W.F., Duguet, M., Gaasterland, T., Garrett, R.A., Ragan, M.A., Sensen, C.W., Van der Oost, J. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 7835-7840.

- 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. USA*, **71**, 1342-1346.
- Si, K. and Maitra, D. (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., Yamamoto, Y. and Asano, K. (2004) Physical association of eukaryotic initiation factor 5 (eIF5) carboxyl terminal domain with the lysine-rich eIF2 $\beta$  segment strongly enhances its binding to eIF3. *J. Biol. Chem.*, **279**, 49644-49655.
- Szkaradkiewicz, K., Zuleeg, T., Limmer, S. and Sprinzl, M. (2000) Interaction of fMet-tRNA<sup>Met</sup> and fMet-AMP with the C-terminal domain of *Thermus thermophilus* translation initiation factor 2. *Eur. J. Biochem.*, **267**, 4290-4299.
- Sonenberg, N. (1988) Cap-binding proteins of eukaryotic messenger RNA: functions in initiation and control of translation. *Prog. Nucleic Acid Res. Mol. Biol.*, **35**, 173-207.
- Sonenberg, N., Hershey, J.W.B. and Mathews, M.B. (2000) Translational control of gene expression. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Sonenberg, N. and Dever, T.E. (2003) Eukaryotic translation initiation factors and regulators. *Current Opinion in Structural Biology*, **13**, 56-63.
- Sørensen, H.P., Hedegaard, J., Sperling-Petersen, H.U. and Mortensen, K.K. (2001) Remarkable conservation of translation initiation factors: IF1/eIF1A and IF2/eIF5B are universally distributed phylogenetic markers. *IUBMB Life*, **51**, 321-327.
- Spahn, C.M.T., 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.
- Sprang, S.R. (1997) G PROTEIN MECHANISMS: Insights from structural analysis. *Annu. Rev. Biochem.*, **66**, 639-678.
- Sprengart, M.L., Fatscher, H.P. and Fuchs, E. (1990) The initiation of translation in *E. coli*: apparent base pairing between the 16s rRNA and downstream sequences of the mRNA. *Nucleic Acids Res.*, **18**, 1719-1723.
- 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.
- Sundari, R.M., Stringer, E.A., Shulman, L.H. and Maitra, U. (1976) Interaction of bacterial initiation factor 2 with tRNA. *J. Biol. Chem.*, **251**, 3338-3345.
- Tahara, M., Ohsawa, A., Saito, S. and Kimura, M. (2004) *In vitro* phosphorylation of initiation factor 2 $\alpha$  (eIF2 $\alpha$ ) from hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *J. Biol. Chem.*, **135**, 479-485.
- Terwilliger, T.C. and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D55*, 849-861.
- Terwilliger, T.C. (1999) Reciprocal-space solvent flattening. *Acta Crystallogr. D55*, 1863-1871.
- Thompson, G.M., Pacheco, E., Melo, E.O. and Castilho, B.A. (2000) Conserved sequences in the  $\beta$  subunit of archaeal and eukaryal translation initiation factor 2 (eIF5), absent from eIF5, mediate interaction with eIF2 $\gamma$ . *Biochem. J.*, **347**, 703-709.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876-4882.
- Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Sasaki, M., Taniguchi, T., Miyagawa, H., Kitamura, K., Miura, K. and Ishida, T. (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.
- Unbehauen, A., Borukhov, S.I., Hellen, C.T.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., 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., Mathew, A.M., 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.
- Vasquez 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. USA*, **90**, 7215-7219.

- Vetter, I.R. and Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science*, **294**, 1299-1304.
- Vitagliano, L., Masullo, M., Sica, F., Zagari, A. and Bocchini, V. (2001) The crystal structure of *Sulfolobus solfataricus* elongation factor 1alpha in complex with GDP reveals novel features in nucleotide binding and exchange. *EMBO J.*, **20**, 5305-5311.
- 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.
- Wakao, H., Romby, P., Westhof, E., Laalami, S., Grunberg-Manago, M., Ebel, J.P., Ehresmann, C. and Ehresmann, B. (1989) The solution structure of the *Escherichia coli* initiator tRNA and its interactions with initiation factor 2 and the ribosomal 30 S subunit. *J. Biol. Chem.*, **264**, 20363-20371.
- Wakao, H., Romby, P., Laalami, S., Ebel, J.P., Ehresmann, C., and Ehresmann, B. (1990) Binding of initiation factor 2 and initiator tRNA to the *Escherichia coli* 30S ribosomal subunit induces allosteric transitions in 16S rRNA. *Biochemistry*, **29**, 8144-8151.
- Wang, Y., Jiang, Y., Meyering-Voss, M., Sprinzl, M. and Sigler, P.B. (1997) Crystal structure of the EF-Tu·EF-Ts complex from *Thermus thermophilus*. *Nat. Struct. Biol.*, **4**, 650-656.
- 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.B. (1982) The binding of fluorescein-labeled protein synthesis initiation factor 2 to *Escherichia coli* 30S ribosomal subunits determined by fluorescence polarization. *J. Biol. Chem.*, **257**, 1215-1220.
- 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.
- Westermann, P., Nygard, O. and Bielka, H. (1980) The  $\alpha$  et  $\gamma$  subunits of initiation factor eIF2 can be cross-linked to 18S ribosomal RNA within the quaternary initiation complex eIF2•Met-tRNA<sup>f</sup>•GDPCP•small ribosomal subunit. *Nucleic Acids Res.*, **8**, 3065-3071.
- Williams, D.D., Price, N.T., Loughlin, A.J. and Proud, C.G. (2001) Characterization of the mammalian initiation factor eIF2B complex as a GDP dissociation stimulator protein. *J. Biol. Chem.*, **276**, 24697-24703.
- Williams, N.P., Hinnebusch, A.G. and Donahue, T.F. (1989) Mutations in the structural genes for eukaryotic initiation factors 2 alpha and 2 beta of *Saccharomyces cerevisiae* disrupt translational control of GCN4 mRNA. *Proc. Natl. Acad. Sci. USA*, **86**, 7515-7519.
- Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., Morgan-Warren, R.J., Carter, A.P., Vornrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327-339.

- Woese, C.R. and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. USA*, **74**, 5088-5090.
- Woese, C.R., Mgarum, L.J. and Fox, G.E. (1978) Archaeobacteria. *J. Mol. Evol.*, **11**, 245-251.
- 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. USA*, **87**, 4576-4579.
- 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.
- Xiong, Y. and Steitz, T.A. (2004) Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature*, **430**, 640-645.
- 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 Å resolution. *Science*, **292**, 883-896.
- Yusupova, G., Reinbolt, J., Wakao, H., Laalami, S., Grunberg-Manago, M., Romby, P., Ehresmann, B. and Ehresmann, C. (1996) Topography of the *Escherichia coli* initiation factor 2/fMet-tRNA(f)(Met) complex as studied by cross-linking. *Biochemistry*, **35**, 2978-2984.
- 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.
- Zoll, W.L., Horton, L.E., Komar, A.A., Hensold, J.O. and Merrick, W.C. (2002) Characterization of mammalian eIF2A and identification of the yeast homolog. *J. Biol. Chem.*, **277**, 37079-37087.

# *Liste des Travaux*



### **Publications:**

Yatime, L., Mechulam, Y., Blanquet, S. and Schmitt, E. (2005) Structural switch of the  $\gamma$  subunit in an archaeal aIF2 $\alpha\gamma$  dimer. *Structure*, in press.

Yatime, L., Schmitt, E., Blanquet, S. and Mechulam, Y. (2005) Structure-function relationships of the intact aIF2 $\alpha$  subunit from the archaeon *Pyrococcus abyssi*. *Biochemistry*, **44**, 8749-8756.

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.

### **Communications orales:**

Yatime, L., Schmitt, E., Blanquet, S. and Mechulam, Y. Rôle du facteur d'initiation aIF2 dans le démarrage de la traduction chez les Archées, 5<sup>ème</sup> Rencontre sifrARN, 10-13 octobre 2004, Arcachon, France.

Yatime, L., Schmitt, E., Blanquet, S. and Mechulam, Y. Structural and functional studies of archaeal initiation factors, *FP6 Meeting*, 10-12 octobre 2003, Aarhus, Danemark.

Yatime, L., Schmitt, E., Blanquet, S. and Mechulam, Y. Functional dissection of the heterotrimeric archaeal initiation factor aIF2, 20<sup>th</sup> *International tRNA Workshop*, 2-7 octobre 2003, Banz, Allemagne.

### **Posters:**

Yatime, L., Schmitt, E., Blanquet, S. and Mechulam, Y. Structure-function relationships of the intact aIF2 $\alpha$  subunit from the archaeon *Pyrococcus abyssi*, EMBO Conference on Protein Synthesis and Translation Control, EMBL Heidelberg, Germany, September 14-18, 2005.

