



**HAL**  
open science

# Mécanisme de formation du complexe de démarrage de la traduction chez les Archées

Auriane Monestier

► **To cite this version:**

Auriane Monestier. Mécanisme de formation du complexe de démarrage de la traduction chez les Archées. Biochimie, Biologie Moléculaire. Université Paris Saclay (COMUE), 2016. Français. NNT : 2016SACLX047 . tel-01513313

**HAL Id: tel-01513313**

**<https://pastel.hal.science/tel-01513313>**

Submitted on 24 Apr 2017

**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.

NNT : 2016SACLX047

THESE DE DOCTORAT  
DE  
L'UNIVERSITE PARIS-SACLAY  
PREPAREE A  
L'ECOLE POLYTECHNIQUE

ECOLE DOCTORALE N° 573  
INTERFACES

Spécialité de doctorat : Biologie

Par

**Auriane Monestier**

Etude du complexe de démarrage de la traduction chez les Archées

**Thèse présentée et soutenue à l'École Polytechnique, le 14 octobre 2016**

**Composition du Jury :**

Dr Carine Tisne	Directrice de recherche, Université Paris Descartes	Rapporteur
Dr Pascale Romby	Directrice de recherche, IBMC	Rapporteur
Dr Soumaya Laalami	Chargée de recherche, IBPC	Examineur
Pr Yves Méchulam	Directeur de recherche, BIOC	Président du jury
Dr Emmanuelle Schmitt	Directrice de recherche, BIOC	Directrice de thèse



## Remerciements

Tout d'abord, je voudrais remercier Pascale Romby et Carine Tisne d'avoir accepté de relire cette thèse et d'en être rapporteurs. Je remercie également Soumaya Laalami d'avoir accepté d'être examinatrice de ce travail. Je remercie les membres du jury d'avoir accepté d'assister à la présentation de ce travail.

Ce travail a été réalisé à l'Ecole Polytechnique. Je tiens grandement à remercier Yves Mechulam, directeur du laboratoire et président du département de biologie, pour m'avoir offert l'opportunité d'accomplir ce travail au sein de son équipe. En particulier pour avoir distribué ma demande de stage au sein du laboratoire il y a 7 ans et d'avoir accepté mon retour chaque année depuis.

Je tiens à exprimer mes plus vifs remerciements à Emmanuelle Schmitt qui a accepté d'encadrer ce travail doctoral, pour ses multiples conseils et pour toutes les heures consacrées à diriger cette recherche. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité et sa patience. J'ai été extrêmement sensible à ses qualités d'écoute et de compréhension tout au long de ces années.

Je voudrais aussi remercier tout particulièrement Pierre-Damien Coureux, qui a recueilli ma demande de stage il y a 7 ans, qui a pris le temps de m'enseigner avec une grande patience les bases du travail en laboratoire, qui m'a supportée dans son bureau toutes ces années et qui a toujours su trouver les mots pour me guider vers la réussite.

Je remercie l'équipe de B2, Etienne Dubiez, Christine Lazennec-Schurdevin, Michel Panvert, pour leur bonne humeur quotidienne, leur soutien et leurs conseils.

Je remercie également tous les membres du laboratoire pour leur sympathie, notamment ceux avec qui j'ai eu l'occasion de travailler et les autres simplement pour les bons moments partagés.

Ces remerciements seraient incomplets si je n'exprimais pas ma reconnaissance à ceux qui ont plus particulièrement assuré le soutien affectif de ce travail doctoral : mes parents, qui ont toujours su me soutenir et m'encourager, mes amis, et plus particulièrement Vincent, qui durant ces trois années de thèse est resté à mes côtés, m'a comprise, soutenue, et m'a toujours exprimé son amour et sa confiance durant les moments de doutes.



## Liste des abréviations

aaRS	aminoacyl-ARNt synthétase
ADN	Acide DésoxyriboNucléique
aIF	Facteur archéen du démarrage de la traduction
Amp	Ampicilline
ARN	Acide RiboNucléique
ARNm	ARN messenger
ARNr	ARN ribosomal
ARNt	ARN de transfert
ARNt <sub>i</sub> <sup>Met</sup>	ARN de transfert initiateur methionine
ARNt <sub>m</sub> <sup>Met</sup>	ARN de transfert élongateur methionine
aSD	Anti Shine-Dalgarno
ATP	Adénosine TriPhosphate
Cm	Chloramphénicol
DO	Densité Optique
EDTA	EthylèneDiamine-tétra-Acétate
eIF	Facteur eucaryotique de démarrage de la traduction
Met-ARNt <sub>i</sub> met	ARNt initiateur bactérien méthionylé et formylé
GAP	GTPase Activating Protein
GDPNP	5'-Guanylylimido-DiPhosphate
Gm	Gentamycine
GTP/GDP	Guanosine Tri/DiPhosphate
HEPES	Acide 4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonique
IF1, IF2, IF3	Facteurs bactériens du démarrage de la traduction 1, 2 et 3
IPTG	IsoPropyl-b-D-ThioGalactoside
Kan	Kanamycine
Kd	Constante de dissociation
Met-ARNt <sub>i</sub> met	ARNt initiateur eucaryotique ou archéen méthionylé
MetRS	méthionyl-ARNt synthetase
MTF	methionyl-ARNt <sub>i</sub> <sup>Met</sup> formyltransférerase
OB fold	Oligonucleotide/Oligosaccharide Binding fold
PABP	Poly(A) Binding Protein
PCR	Polymerase Chain Reaction

PMSF	PhenylMethylSulfonyl Fluoride
SD	Shine-Dalgarno
SDS	DodécylSulfate de Sodium
SDS-PAGE	Electrophorèse sur Gel de PolyAcrylamide en présence de SDS
Tris	Tris(hydroxyméthyl)aminométhane
5'-UTR	Région 5' non traduite (UnTranslated Region)

## Table des matières

<b>1 Introduction .....</b>	<b>13</b>
<b>1.1 Les différents acteurs du démarrage de la traduction .....</b>	<b>15</b>
1.1.1 Les ARNts initiateurs.....	15
1.1.1.1 L'ARNt <sub>f</sub> méthionine bactérien .....	16
1.1.1.2 Les ARNt <sub>i</sub> méthionine eucaryote et archée.....	20
1.1.2 Le ribosome .....	21
1.1.2.1 Description générale .....	21
1.1.2.2 Structure générale du ribosome et sites de liaison à l'ARNt.....	22
1.1.2.3 Fixation de l'ARNt initiateur au site P du ribosome .....	23
1.1.3 Les facteurs de démarrage de la traduction.....	24
<b>1.2 Le démarrage de la traduction dans les différents domaines du vivant .....</b>	<b>26</b>
1.2.1 Le démarrage bactérien .....	26
1.2.2 Le démarrage eucaryote.....	30
1.2.3 Le démarrage archée.....	33
<b>1.3 Le complexe cœur du démarrage de la traduction .....</b>	<b>36</b>
1.3.1 Le facteur hétérotrimérique e/αIF2 .....	36
1.3.1.1 Identification du rôle d'eIF2 dans la sélection du codon de démarrage.....	36
1.3.1.2 Le complexe ternaire (CT) e/αIF2-GTP-ARNt <sub>i</sub> Met .....	37
1.3.2 Les facteurs IF1 et α/eIF1A .....	38
1.3.2.1 Le facteur IF1 .....	38
1.3.2.2 Structures des facteurs α/eIF1A .....	41
1.3.2.2.1 Le facteur eucaryote eIF1A.....	41
1.3.2.2.2 Le facteur archée αIF1A.....	42
1.3.3 Structure des facteurs IF3 et α/eIF1 .....	44
1.3.3.1 Le facteur bactérien IF3.....	44
1.3.3.2 Structure des facteurs eucaryotes et archées e/αIF1.....	46
1.3.3.2.1 Le facteur eucaryote eIF1.....	46
1.3.3.2.2 Le facteur archée αIF1 .....	47
1.3.4 Fonction des facteurs eucaryotes eIF1 et eIF1A .....	48
1.3.4.1 Rôle des facteurs eIF1 et eIF1A dans le balayage de l'ARNm par le complexe de démarrage de la traduction.....	48

1.3.4.2 Les différentes régions d'eIF1A nécessaires à la reconnaissance du bon codon de démarrage de la traduction.....	49
1.3.4.3 Le départ d'eIF1 suite à l'appariement codon/anticodon.....	50
<b>1.4 Travail de thèse.....</b>	<b>51</b>
1.4.1 Construction d'une souche d'E. coli permettant la production d'ARNt initiateur d'intérêt .....	52
1.4.2 Etude structurale du complexe de démarrage archée .....	53
1.4.3 Etude du rôle du facteur aIF1 lors du démarrage de la traduction .....	53
<b>2 Etude structurale de l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>.....</b>	<b>55</b>
<b>2.1 Introduction.....</b>	<b>56</b>
<b>2.2 Discussion de l'étude de l'ARNt<sub>f</sub>MetA1-U72 .....</b>	<b>84</b>
<b>3 Etude du rôle d'aIF1 dans le complexe de démarrage de la traduction ...</b>	<b>89</b>
<b>3.1 Etude de la liaison du facteur aIF1 sur la sous-unité du ribosome par anisotropie de fluorescence.....</b>	<b>93</b>
3.1.1 Obtention des acteurs du complexe de démarrage et principe de l'anisotropie de fluorescence.....	93
3.1.1.1 Obtention de petites sous-unités ribosomales 30S .....	93
3.1.1.2 Purification du facteur de démarrage aIF1 et de ses variants .....	94
3.1.2 L'anisotropie de fluorescence : principe général .....	95
3.1.2.1 Marquage des protéines .....	97
3.1.3 Mise au point du titrage d'aIF1 par la petite sous-unité 30S du ribosome .....	98
3.1.3.1 Titrage de l'ARNm modèle avec la sous-unité 30S du ribosome .....	98
3.1.3.2 Titrage du facteur aIF1 de P. abyssi marqué sur la cystéine 63 avec la petite sous-unité 30S du ribosome et l'ARNm.....	99
3.1.3.3 Titrage du facteur aIF1* marqué sur la cystéine 46 et mise au point de la mesure de la constante de dissociation du complexe 30S:ARNm:aIF1 par anisotropie de fluorescence	101
3.1.3.4 Etude de la liaison d'aIF1* à la sous-unité 30S seule .....	105
3.1.4 Caractérisation des résidus d'aIF1 impliqués dans la liaison au complexe 30S:ARNm ..	106
3.1.4.1 Rôle de la boucle 30-35 dans la liaison d'aIF1 au complexe 30S:ARNm .....	106
3.1.4.1.1 Caractérisation du mutant aIF1*-polyAla.....	106
3.1.4.1.2 Caractérisation du mutant aIF1*-YA <sub>33</sub> .....	107
3.1.4.1.3 Caractérisation du mutant aIF1*-KA <sub>35</sub> .....	108
3.1.4.1.4 Caractérisation du mutant aIF1*-RA <sub>32</sub> .....	108
3.1.4.2 Rôle de l'extrémité N-terminale dans la liaison au complexe 30S:ARNm : caractérisation du mutant aIF1*-ΔN .....	109

<b>3.2 Etude de la formation et de la stabilité du complexe de démarrage sur l'ARNm</b> .....	<b>111</b>
3.2.1 Principe du « toeprint » .....	111
3.2.2 Conditions de stabilisation de la sous-unité 30S du ribosome sur l'ARNm.....	115
3.2.2.1 Stabilisation du ribosome sur l'ARNm.....	115
3.2.2.2 Caractérisation de la saturation du ribosome en facteur aIF1.....	117
3.2.2.3 Influence de la séquence de l'ARNm dans la formation du complexe aIF1:30S sur l'ARNm... ..	118
3.2.2.3.1 Influence de la position du codon de démarrage sur l'ARNm.....	118
3.2.2.3.2 Influence de la séquence SD.....	121
3.2.2.4 Liaison de l'ARNt initiateur au complexe 30S: ARNm .....	123
3.2.3 Etude de différent complexes de démarrage sur la stabilité du complexe 30S:ARNm... ..	125
3.2.3.1 Le complexe CT:30S:ARNm.....	125
3.2.3.2 Variation de la concentration en ARNt.....	127
3.2.3.3 Le complexe Met-ARNtA <sub>1</sub> U <sub>72</sub> :aIF2:aIF1:30S:ARNm (CI sans aIF1A) .....	127
3.2.3.4 Signal de toeprint du complexe de démarrage complet : Met-ARNtA <sub>1</sub> U <sub>72</sub> :aIF2:aIF1:aIF1A:30S:ARNm (CI) .....	128
3.2.4 Etude du rôle des résidus de la boucle basique d'aIF1 .....	130
3.2.4.1 Etude du rôle de l'arginine 32 .....	131
3.2.4.1.1 Le complexe aIF1-RA32:30S sur l'ARNm WT .....	131
3.2.4.1.2 Le CI avec aIF1-RA32.....	132
3.2.4.2 Etude du rôle de la tyrosine 33 .....	134
3.2.4.2.1 Le complexe aIF1-YA33:30S sur l'ARNm WT : rôle dans la stabilisation de la 30S sur l'ARNm..... ..	134
3.2.4.2.2 Le CI avec aIF1-YA33.....	135
3.2.4.3 Etude du rôle de la lysine 35 .....	137
3.2.4.4 Etude du double-mutant aIF1-RA32-YA33 .....	138
3.2.4.4.1 Le complexe aIF1-RA32-YA33:30S sur l'ARNm WT : rôle dans la stabilisation de la 30S sur l'ARNm .....	139
3.2.4.4.2 Etude du variant aIF1-RA32-YA33 dans le CI.....	140
3.2.4.5 Etude du double-mutant aIF1-RA32-KA35 .....	142
3.2.4.5.1 Le complexe aIF1-RA32-KA35:30S sur l'ARNm WT : rôle dans la stabilisation de la 30S sur l'ARNm .....	142
3.2.4.5.2 Etude du variant aIF1-RA32-KA35 dans le CI.....	143
3.2.5 Etude du rôle du domaine N-terminal d'aIF1 dans la formation du complexe de démarrage de la traduction .....	144
3.2.5.1 Etude du complexe aIF1-ΔN:30S:ARNm .....	144

3.2.5.2 Etude de l'extrémité N-terminale dans le CI .....	145
3.2.6 Rôle d'aIF1 dans le rejet d'un mauvais complexe de démarrage .....	146
3.2.6.1 Utilisation d'un ARNt élongateur dans la formation du complexe de démarrage de la traduction .....	146
3.2.6.2 Etude de l'influence des 3 paires de bases GC de l'ARNti dans la formation du complexe de démarrage de la traduction sur l'ARNm .....	148
<b>4 Etude structurale du complexe de démarrage de la traduction archée ...</b>	<b>151</b>
<b>5 Discussion et perspectives .....</b>	<b>193</b>
5.1 Discussion .....	194
5.2 Perspectives .....	199
<b>6 Matériels et méthodes .....</b>	<b>203</b>
6.1 Inactivation de gènes chez E.coli par la technique lambda red .....	204
6.1.1 Souches bactériennes et plasmides .....	204
6.1.2 Préparation des produits de PCR possédant le gène KanR, 2 séquences FRT et 2 régions homologues au(x) gène(s) à inactiver .....	205
6.1.3 Inactivation du gène metY (inactivation primaire).....	206
6.1.4 Elimination du marqueur Kan puis élimination du plasmide pLH29 .....	207
6.1.5 Test de complémentation fonctionnel.....	207
6.1.6 Inactivation du gène metY (inactivation secondaire).....	208
6.2 Purification des sous-unités 30S .....	209
6.3 Expression et purification des différents mutants d'aIF1.....	210
6.4 Marquage des différents mutants d'aIF1 avec la coumarine .....	212
6.5 L'anisotropie de fluorescence .....	213
6.6 L'expérience de toeprint.....	215
6.7 Isolation de petits ARNts .....	216
<b>7 Bibliographie .....</b>	<b>219</b>

## **Avant-propos**

La première phase du décodage de l'information génétique est appelée transcription. Au cours de cette étape, de petites portions du génome (ADN) sont transcrites en fragments d'acides ribonucléiques messagers (ARNm). Ceux-ci sont composés, tout comme l'ADN, d'un code à 4 lettres à la différence qu'un atome d'oxygène supplémentaire est présent sur chacun des nucléotides de l'ARN. Le ribose remplace alors le désoxyribose. Cet atome d'oxygène supplémentaire confère une plus grande flexibilité aux molécules d'ARN par rapport aux molécules d'ADN. Chez les Eucaryotes, les ARNm contiennent un gène unique et sont structurés en mosaïques de section codantes (exons) et non codantes (introns). Ces ARNm primaires devront subir une maturation correspondant à l'excision des introns et la jonction des exons (épissage) ainsi qu'à l'ajout d'une coiffe guanosine triphosphate en 5' et d'une queue poly-adénine en 3'. Chez les bactéries, les ARNm sont polycistroniques, c'est-à-dire qu'ils contiennent plusieurs gènes permettant ainsi la synthèse de plusieurs protéines à partir d'un même ARNm. Ces ARNm contiennent donc plusieurs codons de démarrage et codons d'arrêt.

La seconde phase du décodage de l'information génétique est l'étape de traduction des ARNm en protéines, ou biosynthèse des protéines. Elle permet de passer d'un code à 4 lettres, formé des bases des acides nucléiques (A, U, G, C), à un code à 20 lettres correspondant aux acides aminés entrant dans la composition des protéines.

La traduction est un processus qui se déroule en trois étapes principales, le démarrage, l'allongement et la terminaison.

La correspondance entre l'information génétique portée par les ARNm et les acides aminés contenus dans la protéine est assurée grâce à de petites molécules polynucléotidiques, les ARN de transfert. Ces courts ARN non codant possèdent des extrémités spécifiques constantes dont l'une est une séquence de trois nucléotides appelée anticodon, complémentaire de la séquence d'un triplet de nucléotides de l'ARNm (codon) spécifiant un acide aminé. Les ARN de transfert sont reconnus par des enzymes spécifiques cruciales à la biosynthèse des protéines, les aminoacyl-ARNt-synthétases (aaRS). Les aaRS sont responsables de l'amination spécifique des ARNt par l'acide aminé qui correspond à la séquence de leur anticodon. Au cours du décodage, l'anticodon vient s'apparier au codon complémentaire sur l'ARNm, permettant ainsi la correspondance entre le codon et l'acide aminé estérifié à l'extrémité 3' de l'ARNt (voir figure ci-dessous). Celui-ci sera alors attaché au polypeptide naissant. 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.

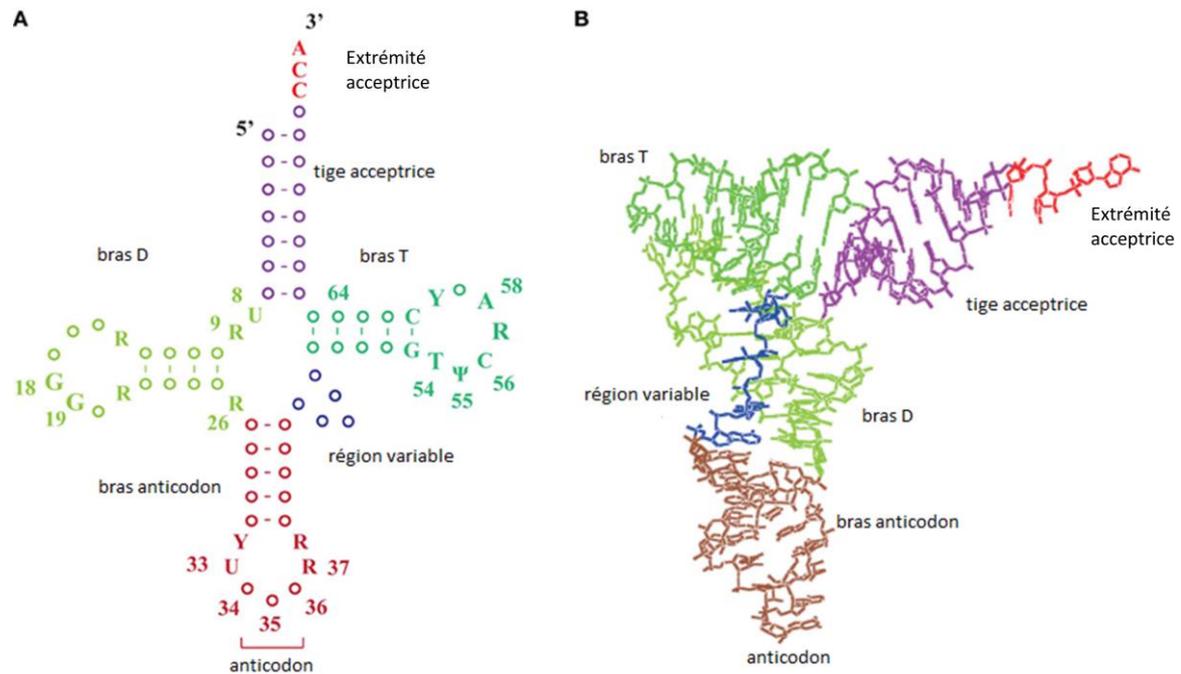


Figure 0 : A/ Structure secondaire en feuille de trèfle d'un ARNt B/ Structure tertiaire en L d'un ARNt  
D'après Hori 2014

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. Un décalage d'un nucléotide entraînerait la synthèse d'un polypeptide non fonctionnel, voir toxique pour la cellule, avec une chaîne aberrante d'acides aminés. Il s'agit donc d'une étape majeure de l'ensemble du processus de décodage.

De nombreuses études récentes ont en effet montré que le démarrage de la traduction était une étape critique pour le contrôle de nombreux processus cellulaires comme par exemple la réponse à la privation en nutriments, la réponse synaptique (Clemens 1990)(Costa-Mattioli et al. 2007) ou encore la cancérisation (Clemens et Bommer 1999). La compréhension des mécanismes moléculaires qui gouvernent cette étape constitue donc également un enjeu de par ses applications thérapeutiques possibles (Borck et al. 2012).

Dans cette introduction, après avoir présenté les différents acteurs du démarrage de la traduction, je décrirai les différents mécanismes de démarrage dans les trois domaines. Finalement, nous verrons

que malgré les spécificités de chaque domaine du vivant, il existe chez les eucaryotes et les archées un complexe cœur formé de l'ARNm lié à la petite sous-unité du ribosome en présence de trois protéines : e/aIF1, e/aIF1A, e/aIF2 et de l'ARNt initiateur méthionylé.

Dans une dernière partie, je présenterai donc les connaissances disponibles sur la fonction de ces trois protéines dans le démarrage de la traduction eucaryote et archée ainsi que les objectifs de mon travail de thèse.



# 1 Introduction

L'apparition du microscope électronique vers le milieu du XX<sup>ème</sup> siècle a abouti à la mise en place d'un monde du vivant divisé en deux grands groupes basé sur la structure de leurs cellules : les Eucaryotes, formés de cellules à noyaux, et les Procaryotes, formés de cellules sans noyau.

D'apparence souvent semblable aux bactéries, les archées ont longtemps été considérées comme des bactéries extrémophiles. Les recherches en phylogénétique sur les organismes procaryotes à la fin des années 1970 ont abouti à la publication d'un arbre phylogénétique fondé sur la séquence des gènes des ARN ribosomiques 16S, dans lequel le domaine procaryote est finalement scindé en deux groupes distincts : les Bactéries et les Archées (Woese et Fox 1977) (voir figure 1.1 ci-dessous).

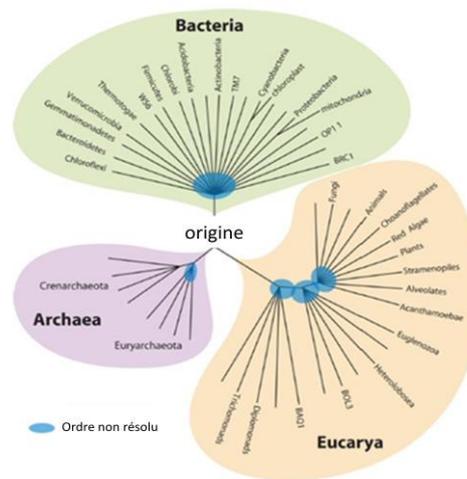


Figure 1.1 : représentation de l'arbre du vivant d'après Kyrpides et Woese 1998

Ces analyses ont montré que les séquences des ARNr 16S d'archées étaient pratiquement aussi éloignées des séquences des ARNr 16S bactériens que de celles des ARNr 18S eucaryotes.

Bien qu'elles possèdent des voies métaboliques qui leurs sont propres (comme la méthanogénèse chez certaines archées), certains mécanismes de survie cellulaire archéens disposent de similitudes avec chacun des deux autres domaines.

Tout comme chez les bactéries, le chromosome archéen est circulaire et ne possède qu'une origine de réplication. Les archées peuvent également posséder des plasmides et leurs gènes ne contiennent pas d'introns. La réplication de l'ADN ainsi que la transcription de l'ADN en ARNm chez les archées sont des étapes réalisées par des enzymes apparentées à celles des eucaryotes et non à celles des bactéries. De même, la présence d'histones dans le matériel génétique des archées rapproche ces dernières des eucaryotes et les distingue des bactéries.

De manière intéressante, on remarque que dans l'arbre du vivant, les protéines impliquées dans la traduction sont les plus conservées. De façon universelle, le démarrage de la traduction nécessite la présence d'un complexe macromoléculaire spécifique impliquant la petite sous-unité du ribosome, un ARN de transfert (ARNt) particulier, appelé ARNt initiateur, ainsi que plusieurs protéines appelées facteurs de démarrage. Ce complexe est centré sur le codon de démarrage AUG qui définit le cadre de lecture de la séquence de l'ARNm. Toute la difficulté de cette étape est donc la reconnaissance du codon de démarrage (AUG le plus souvent), codant systématiquement pour une méthionine initiatrice. Ce codon est spécifiquement reconnu par un ARNt initiateur méthionylé.

Tableau récapitulatif des acteurs clés du démarrage de la traduction :

Acteurs clés du démarrage de la traduction présents dans les trois domaines du vivant
ARNt initiateur
ARNm
Ribosome
Facteurs protéiques appelés facteurs de démarrage

## 1.1 Les différents acteurs du démarrage de la traduction

### 1.1.1 Les ARNts initiateurs

Dans chaque domaine du vivant, il existe deux types d'ARNt aminoacylable par de la méthionine grâce à l'action d'une méthionyl-ARNt-synthétase. Cependant, alors qu'un type sera utilisé pour le décodage des codons internes à la séquence des gènes ( $\text{ARNt}_m^{\text{Met}}$ ), l'autre sera utilisé uniquement lors du décodage du codon méthionine de démarrage de la traduction ( $\text{ARNt}_i^{\text{Met}}$ ). Ce dernier, appelé ARNt initiateur méthionylé est reconnu spécifiquement par les facteurs de démarrage de la traduction, à la différence du premier, qui lui est reconnu par les facteurs d'élongation. La présence de deux types d'ARNt<sup>Met</sup> pourrait permettre d'éviter la concurrence entre les mécanismes de démarrage et d'allongement (Guillon et al. 1996).

La sélection de l'ARNt initiateur au niveau du site P du ribosome par la machinerie de démarrage est donc une étape clé permettant de définir la phase de lecture correcte de l'ARNm. Des

caractéristiques spécifiques de cet ARNt initiateur vont permettre sa sélection par la machinerie de démarrage tout l'en excluant de la machinerie d'allongement.

#### 1.1.1.1 L'ARNt<sub>f</sub> méthionine bactérien

Chez les bactéries, il peut exister jusqu'à deux types d'ARNt initiateur. Chez *Escherichia coli* K12 par exemple, il en existe deux, différents seulement d'un nucléotide en position 46, qui peut être soit une 7-méthyl-guanosine (ARNt<sub>f1</sub><sup>Met</sup> majoritaire) soit une adénine (ARNt<sub>f2</sub><sup>Met</sup> minoritaire).

L'aminocacylation des ARNt<sup>Met</sup> par une méthionyl-ARNt synthétase nécessite la formation d'un complexe enzyme-ARNt, gouvernée par l'anticodon CAU de l'ARNt<sup>Met</sup>. Des déterminants nucléotidiques existant dans la tige acceptrice, une adénosine en position 73 ainsi qu'une paire de base G<sub>2</sub>-C<sub>71</sub> et une paire C<sub>3</sub>-G<sub>70</sub> ont également un rôle dans l'efficacité catalytique de l'enzyme.

Tout comme dans les chloroplastes et les mitochondries, chez les bactéries, la méthionine estérifiée en 3' des ARNt initiateurs est systématiquement N-formylée grâce à l'action d'une méthionyl-ARNt<sup>Met</sup> formyltransférase (MTF). La formylation permet de favoriser la sélection de cet ARNt<sup>Met</sup> lors de la phase de démarrage et son exclusion lors de la phase d'allongement de la traduction. (Guillon et al. 1993; Guillon et al. 1992; Guillon et al. 1996).

En effet, le groupe formyl de l'ARNt initiateur joue un rôle important dans la liaison de l'ARNt<sup>Met</sup> au facteur IF2 et participe au rejet de cet ARNt par le facteur d'allongement EF-Tu.

La détermination de la structure cristallographique de la formylase d'*E. coli* en complexe avec l'ARNt initiateur (Schmitt et al. 1998) a permis de comprendre comment cet ARNt était sélectionné par la MTF (voir figure 1.2).

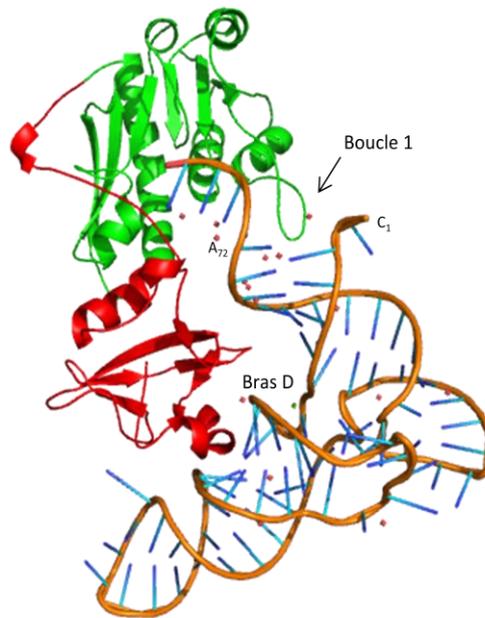


Figure 1.2 : Représentation du complexe formylase fMet-ARNT<sub>f</sub><sup>Met</sup> initiateur d'*E. coli*. L'ARNT est en bleu, le domaine N-terminal de la formylase est en vert et le domaine C-terminal en rouge. D'après Schmitt et al. 1998

Le déterminant d'identité majeur du f-Met-ARNT<sub>f</sub><sup>Met</sup> assurant son interaction spécifique avec la MTF correspond au couple de bases non- appariées C<sub>1</sub>A<sub>72</sub> (Guillon et al. 1992; Seong et RajBhandary 1987) situé au sommet de l'hélice acceptrice de l'ARNT. L'absence d'appariement favorise l'ouverture des bases C<sub>1</sub>A<sub>72</sub> et crée une conformation très particulière sélectionnée par une longue boucle de l'enzyme (boucle L1, voir figure ci après). D'autres déterminants nucléotidiques moins importants interviennent également : A<sub>73</sub>, C<sub>3</sub>G<sub>70</sub> et A<sub>11</sub>U<sub>24</sub> (Lee, Seong, et RajBhandary 1991)

De manière intéressante, il faut souligner que tous les ARNT initiateurs bactériens possèdent un couple de base non appariées à la position 1-72 alors que la plupart des ARNT élongateurs disposent d'une paire de base GC ou CG fortement appariée à cette position.

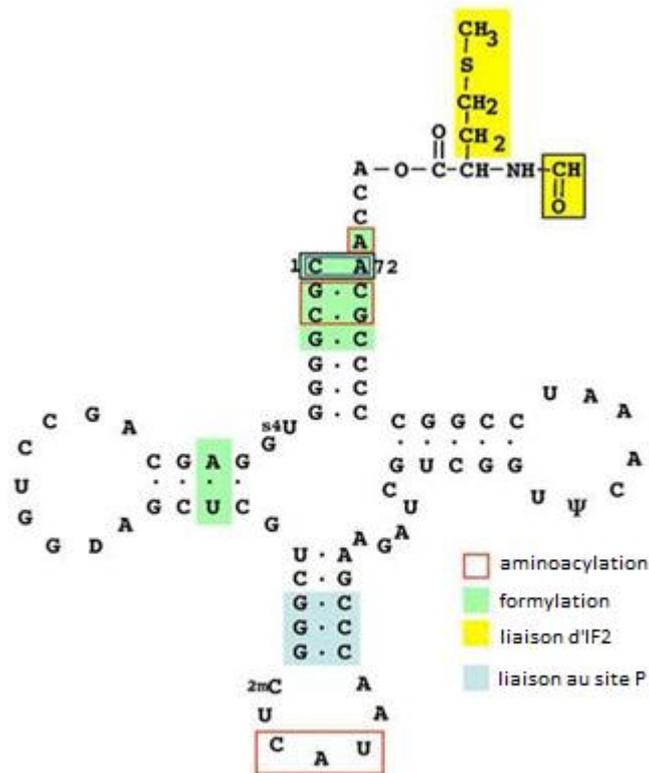


Figure 1.3 : Représentation d'un ARNt<sub>f</sub><sup>Met</sup> initiateur bactérien et rôles des différents déterminants dans le démarrage de la traduction. D'après Mechulam, Blanquet, et Schmitt 2011

En dehors des déterminants nucléotidiques impliqués dans l'étape de formylation, on note également un autre groupe de bases très important pour la spécificité des ARNt initiateurs : les 3 paires de bases consécutives GC situées dans la tige anticodon (figure 1.3 ci-dessus) (Seong et RajBhandary 1987).

De manière intéressante, la structure cristallographique de l'ARNt<sub>f</sub><sup>Met</sup> initiateur d'*E. coli* montre une conformation particulière de la boucle anticodon associée à ces trois paires de bases GC. Cette conformation est liée à une interaction triple entre la base A<sub>37</sub> et les bases G<sub>29</sub>C<sub>41</sub>, un empilement inhabituel de la base A<sub>38</sub> directement avec la base A<sub>36</sub>, ainsi qu'un appariement de type wobble entre les bases C<sub>32</sub> et A<sub>38</sub> (voir figure 5).

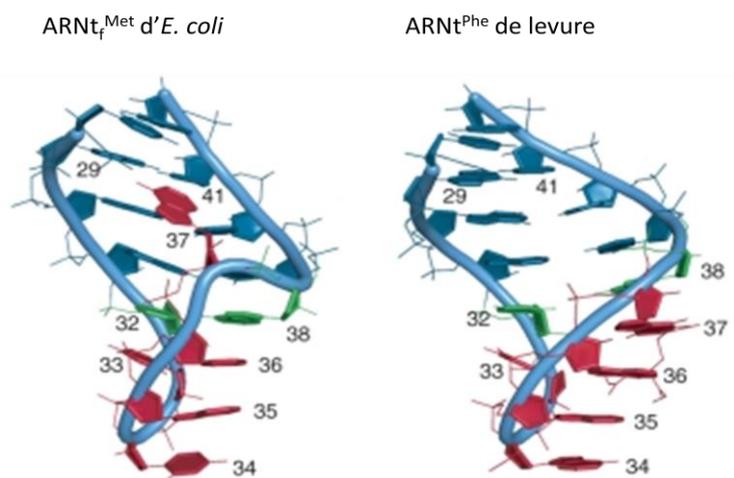


Figure 1.4 : Comparaison entre la structure de la boucle anticodon d'un ARNt initiateur et un ARNt élongateur.

D'après Barraud et al. 2008

Cette conformation se différencie de toutes les structures de boucle anticodon d'ARNt élongateurs connues (Kim et al. 1973; Robertus et al. 1974; Suddath et al. 1974; Jovine, Djordjevic, et Rhodes 2000; Shi et Moore 2000; Westhof, Dumas, et Moras 1985) (voir figure 1.4) et pourrait être importante pour la sélection des ARNt initiateurs au site P de la petite sous-unité ribosomale au moment du démarrage de la traduction.

Malgré tout, la structure de l'ARNt<sup>Met</sup> accommodée au codon de démarrage au niveau du site P du ribosome de *Thermus thermophilus* (Selmer et al. 2006) montre que la conformation classique de la boucle anticodon est restaurée, avec la base A<sub>37</sub> empilée entre les bases U<sub>36</sub> et A<sub>38</sub>. De plus, cette même conformation classique a été retrouvée dans la boucle anticodon d'un ARNt<sup>Met</sup> apparié à l'ARNm au niveau du site P du ribosome (Berk et al. 2006). Ainsi, il ne peut pas être totalement exclu que la conformation inhabituelle observée dans le cristal soit due à l'empilement cristallin.

Cependant, l'analyse de ces différentes structures, sur lesquelles nous reviendrons dans une prochaine partie, laissent entrevoir que cette conformation spécifique de la boucle anticodon de l'ARNt<sup>Met</sup> pourrait permettre à celui-ci d'être reconnu spécifiquement par le complexe de démarrage et pourrait faciliter sa différenciation des ARNt élongateurs. Après l'accommodation au site P du ribosome, la conformation de la boucle anticodon pourrait revenir à une conformation habituelle, observée chez les ARNt élongateurs. L'ARNt<sup>Met</sup> aurait ainsi la capacité de basculer d'une conformation à une autre durant le démarrage de la traduction.

### 1.1.1.2 Les ARN<sub>t</sub><sub>i</sub> méthionine eucaryote et archée

Les ARN<sub>t</sub> initiateurs eucaryotes et archées (ARN<sub>t</sub><sub>i</sub><sup>Met</sup>) ne sont pas formylés. De manière intéressante, ils possèdent de façon quasi systématique une paire de base appariée A<sub>1</sub>U<sub>72</sub> et non plus un couple de base non appariées C<sub>1</sub>A<sub>72</sub> rencontré chez les ARN<sub>t</sub> initiateurs bactériens.

Cette paire de base faible, remplacée par une paire de bases à l'interaction forte G<sub>1</sub>C<sub>72</sub> chez les ARN<sub>t</sub> élongateurs, est aussi un déterminant très important pour la liaison au facteur de démarrage de la traduction archée et eucaryote e/EIF2 et la liaison du complexe e/EIF2:ARN<sub>t</sub><sub>i</sub><sup>Met</sup> ainsi formé à la petite sous-unité du ribosome (Kapp, Kolitz, et Lorsch 2006).

Le bras anticodon de l'ARN<sub>t</sub> initiateur est quant à lui très conservé dans le monde du vivant et contient toujours un anticodon CAU ainsi que les trois paires de bases consécutives GC dans la tige anticodon (figure 1.5). Les mêmes caractéristiques structurales de la tige anticodon bactérienne se retrouvent alors aussi bien chez les ARN<sub>t</sub> initiateurs archée qu'eucaryote.

Cette fois encore les trois paires de bases GC existant dans la tige anticodon des ARN<sub>t</sub><sub>i</sub><sup>Met</sup> ne se retrouvent pas chez les ARN<sub>t</sub> élongateurs archée et eucaryotes, laissant ainsi apparaître leur rôle sans doute déterminant dans la sélection de l'ARN<sub>t</sub> initiateur, ce dont nous discuterons ultérieurement.

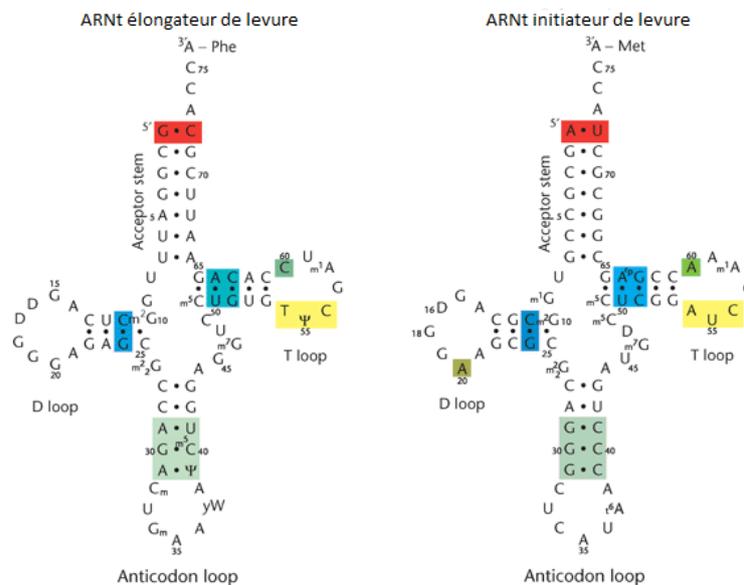


Figure 1.5 : Différences et similitudes entre un ARN<sub>t</sub> élongateur (gauche) et initiateur (droite) de levure.

D'après Marck et Grosjean 2002

Même si les mécanismes de sélection diffèrent selon le domaine du vivant étudié, l'ARNt initiateur possède des caractéristiques qui lui sont propre et qui vont lui permettre de faciliter sa sélection par le système de démarrage et son exclusion par le système d'allongement. Les caractéristiques les plus importantes des ARNts initiateurs sont résumées dans le tableau ci-dessous :

Caractéristiques des ARNt initiateurs selon le domaine du vivant			
	bactéries	archées	eucaryotes
formylation	oui	non	non
bases 1 et 72 appariées	non	paire faible	paire faible
3 paires de bases GC dans la tige anticodon	oui	oui	oui
Facteur impliqué dans la prise en charge de l'ARNt initiateur	IF2	aIF2	eIF2
Séquence de la boucle anticodon CUCAUAA	oui	oui	oui

## 1.1.2 Le ribosome

### 1.1.2.1 Description générale

Le ribosome, un complexe macromoléculaire composé de deux sous-unités, est constitué de trois à quatre molécules d'ARN ribosomiques (ARNr) et de 55 à 80 protéines selon le domaine du vivant. Les ribosomes se retrouvent dans le cytoplasme, liés à la membrane plasmique (chez les bactéries), ou encore liés au niveau de la membrane du réticulum endoplasmique ou de l'enveloppe nucléaire (chez les eucaryotes). 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 (N. Ban et al. 2000; Nissen et al. 2001).

Le ribosome bactérien, tout comme le ribosome archée a un coefficient de sédimentation de 70S. Il est divisé en deux sous-unités :

- la petite sous-unité 30S, contenant l'ARN ribosomiques (ARNr) 16S ainsi que 21 protéines chez *E. coli* et 25 chez *Pyrococcus furiosus*
- la grande sous-unité 50S, contenant les ARNr 23S et 5S, ainsi que 32 protéines chez *E. coli* et 39 chez *P. furiosus*

Le ribosome eucaryote a quant à lui un coefficient de sédimentation plus élevé de 80S.

Il comporte lui aussi deux sous-unités ribosomales :

- la petite sous-unité sédimentant à 40S contient l'ARNr 18S ainsi que 33 protéines ribosomales chez la levure.
- la grande sous-unité sédimentant à 60S contient les ARNr 28S, 58S et 5S ainsi que 46 protéines ribosomales chez la levure.

De manière intéressante, alors qu'aucune protéine ribosomique propre au domaine bactérien ne se retrouve chez les archées, toutes les protéines ribosomiques archées sont communes au domaine eucaryote, ce qui fait du ribosome archée une version miniaturisée de celui des eucaryotes (Nenad Ban et al. 2014).

#### 1.1.2.2 Structure générale du ribosome et sites de liaison à l'ARNt

Les différentes protéines ribosomales ainsi que leur interaction avec les ARNr ribosomaux composant le ribosome ont pu être mises en évidence grâce à des travaux menés avec des sous-unités de ribosome d'*E. coli* 30S (Culver et Noller 1999) ou 50S (Green et Noller 1999) reconstituées à partir de protéines et d'ARNr préalablement isolés et purifiés. Par exemple, le voisinage de la protéine ribosomale L11 dans la sous-unité 50S du ribosome a ainsi pu être décrit et les bases impliquées dans son interaction avec les ARNr ont pu être caractérisées (Holmberg et Noller 1999). De même l'environnement de l'ARNr 5S de la sous-unité 30S a pu être décrit avant la détermination de structures tridimensionnelles (Culver, Heilek, et Noller 1999).

Le rôle crucial du ribosome, son organisation, son mode d'interaction entre les deux sous-unités ainsi que les interactions avec les partenaires ribonucléiques ont pu être mis en avant grâce à la détermination des structures à haute résolution des sous-unités 30S et 50S (N. Ban et al. 2000; Wimberly et al. 2000; Yusupov et al. 2001; Harms et al. 2001). Depuis, la structure du ribosome complet 70S bactérien a été résolue à 3.5 Å de résolution (Schuwirth et al. 2005), celle du ribosome 70S archée à 6.6 Å (Armache et al. 2013) et les structures du ribosome 70S bactérien en complexe

avec l'ARNt et l'ARNm ont également été résolues successivement à 5.5 Å puis à 2.8 Å de résolution (Yusupov et al. 2001; Selmer et al. 2006).

Très récemment, les structures de ribosomes eucaryotes (Ben-Shem et al. 2011; Khatter et al. 2015) ainsi que celle de ribosomes mitochondriaux (Greber et al. 2015) ont également été résolues.

On peut observer les trois sites de liaison pour les molécules d'ARNt : le site A (aminoacyl) où se fixent les ARNt élongateurs, le site P (peptidyl) qui fixe l'ARNt portant la chaîne polypeptidique naissante et accueille l'ARNt initiateur lors de l'étape de démarrage, et le site E (exit) de libération des ARNt déacylés du ribosome (voir figure 1.6)

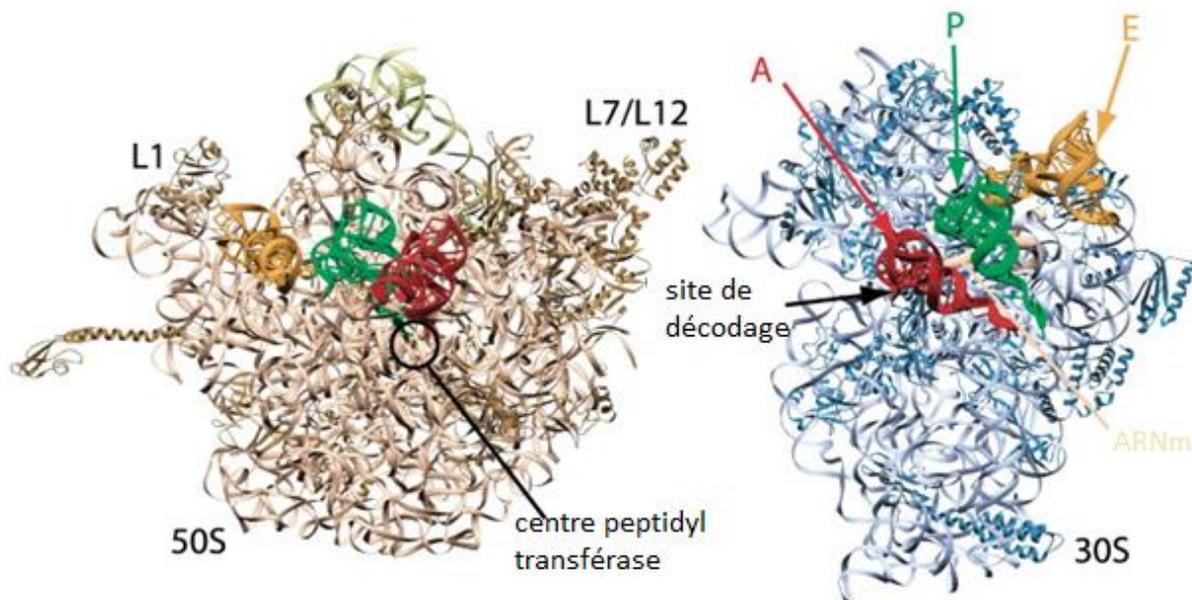


Figure 1.6 : Vue de la sous-unité 50S (gauche) et 30S (droite) du ribosome bactérien. En couleur se trouvent les ARNt aux différents sites de fixation possible A, E, P. D'après Ramakrishnan 2002

### 1.1.2.3 Fixation de l'ARNt initiateur au site P du ribosome

Les avancées structurales ont en particulier permis de montrer les caractéristiques de liaison de l'ARNt initiateur au niveau du site P du ribosome (Selmer et al. 2006). Cette liaison est tout d'abord indépendante de l'interaction avec le codon de l'ARNm (Tomsic et al. 2000).

On observe notamment l'absence d'appariement Watson-Crick des bases C<sub>1</sub>A<sub>72</sub>, la paire de base purine-pyrimidine 11-24 et les trois paires des bases G-C consécutives dans la boucle anticodon. De nombreuses riboprotéines ainsi que des ARN ribosomiaux sont en contact avec l'ARNt initiateur, notamment *via* sa boucle anticodon. Deux des trois paires de bases GC consécutives sont en contact

direct avec l'ARNr 16S. Ces interactions sont plus importantes que si des paires de bases de type A-U (Battle et Doudna 2002) étaient présentes.

Ces observations pourraient expliquer la meilleure stabilisation de l'ARNt initiateur par rapport aux autres ARNt au niveau du site P du ribosome. D'autres contacts peuvent cependant être nécessaires pour assurer la spécificité de liaison de l'ARNt initiateur au site P, qui ne peuvent être révélés par la structure en absence de facteurs de démarrage.

Caractéristiques principales du ribosome selon le domaine du vivant			
	bactérie	archées	eucaryote
Vitesse de sédimentation des sous-unités ribosomales	30S / 50S	30S / 50S	40S/60S
Sites de liaison de l'ARNt au ribosome	Sites A, E, P	Sites A, E, P	Sites A, E, P
ARN ribosomiaux présents	ARNr 23S, 16S, 5S	ARNr 23S, 16S, 5S	28S, 18S, 58S et 5S
Nombre de protéines ribosomales de la petite sous-unité et caractéristiques	21 protéines 7 B 14 BEA	27 protéines 12 EA 15 BEA	31 protéines 4 E 12 EA 15 BEA

B = spécifiques aux bactéries E = spécifiques aux eucaryotes A = spécifiques aux archées

### 1.1.3 Les facteurs de démarrage de la traduction

Pour assurer la fidélité du démarrage de la traduction, des facteurs protéiques, appelés facteurs de démarrage sont nécessaires. La purification de ribosome en présence d'extrait cellulaire et l'analyse des ces extraits sur gradients de sucrose ainsi que sur gel polyacrylamide dénaturant a permis de caractériser un certain nombre de facteurs protéiques nécessaires dans le démarrage de la traduction.

Ceux-ci varient selon le domaine du vivant étudié. Voir tableau ci-dessous :

Les facteurs de démarrage de la traduction		
bactériens	archéens	eucaryotes
<u>IF1 (monomère)</u>	<u>aIF1A (monomère)</u>	<u>eIF1A (monomère)</u>
IF2 (monomère)	aIF5B	eIF5B
	<u>aIF2 (3 sous-unités)</u>	<u>eIF2 (3 sous-unités)</u>
		eIF2A
	aIF2B 3 sous-unités 3 régulatrices : $\alpha$ , $\beta$ , $\delta$	eIF2B 5 sous-unités 3 régulatrices : $\alpha$ , $\beta$ , $\delta$ 2 catalytiques : $\gamma$ , $\epsilon$
<u>IF3</u>	<u>aIF1</u>	<u>eIF1</u>
		eIF3 au moins 6 sous-unités
		eIF4A
		eIF4B
		eIF4E
		eIF4G
		eIF5

L'analyse génomique a permis de mettre en évidence le caractère très conservé de certains de ces facteurs dans les différents domaines du vivant (Kyrpides et Woese 1998). Les homologies observées amènent à supposer des fonctions similaires par ces facteurs dans le mécanisme de démarrage.

Cependant, certains facteurs demeurent spécifiques du domaine eucaryote, résultat direct de l'évolution de ce domaine. Par exemple le complexe eIF4F, comprenant les facteurs eIF4A, E, G et la protéine accessoire eIF4B, est impliqué dans la reconnaissance spécifique de la coiffe des ARNm. Cette maturation des ARNm n'existant que chez les eucaryotes, ce complexe est une évolution eucaryote ne se retrouvant pas dans les autres domaines du vivant. De même le complexe eIF3

impliqué dans la liaison entre les facteurs eIF4 et le complexe de pré-initiation formé des facteurs restant résulte de l'évolution plus complexe du démarrage de la traduction eucaryote. D'autres facteurs, comme le facteur archée et eucaryote a/eIF2, sont quant à eux spécifiques de ces deux domaines et ne retrouvent pas d'homologues bactériens. L'absence de formylation des ARNt méthionylés chez les eucaryotes et chez les archées nécessite un mécanisme de sélection différent exercé par le facteur a/eIF2. Chez les eucaryotes ce mécanisme est accompagné en plus des facteurs eIF2B et eIF5.

Malgré des connaissances croissantes sur les différents acteurs permettant la sélection du codon de démarrage sur l'ARNm, les mécanismes moléculaires mis en jeu restent flous. La fonction et le mécanisme de chacun des facteurs protéiques impliqués continuent de susciter de nombreuses questions.

Avant d'aborder les questions qui ont intéressé mon travail de thèse, je présenterai les grandes lignes des mécanismes de démarrage de la traduction dans les trois domaines du vivant.

## **1.2 Le démarrage de la traduction dans les différents domaines du vivant**

### **1.2.1 Le démarrage bactérien**

Le démarrage de la traduction d'un ARNm bactérien requiert 3 facteurs de démarrage de la traduction (IF) : IF1, IF2 et IF3, un ARNt initiateur chargé avec une formyl-méthionine (ARNt<sub>f</sub><sup>Met</sup> initiateur)(Guillon et al. 1993) et la petite sous-unité du ribosome. Ces 3 facteurs sont essentiels au contrôle des différentes étapes du démarrage, et les liaisons de ces 3 facteurs à la sous-unité 30S sont indépendantes les unes des autres (Milón et al. 2012) (voir figure 1.7)

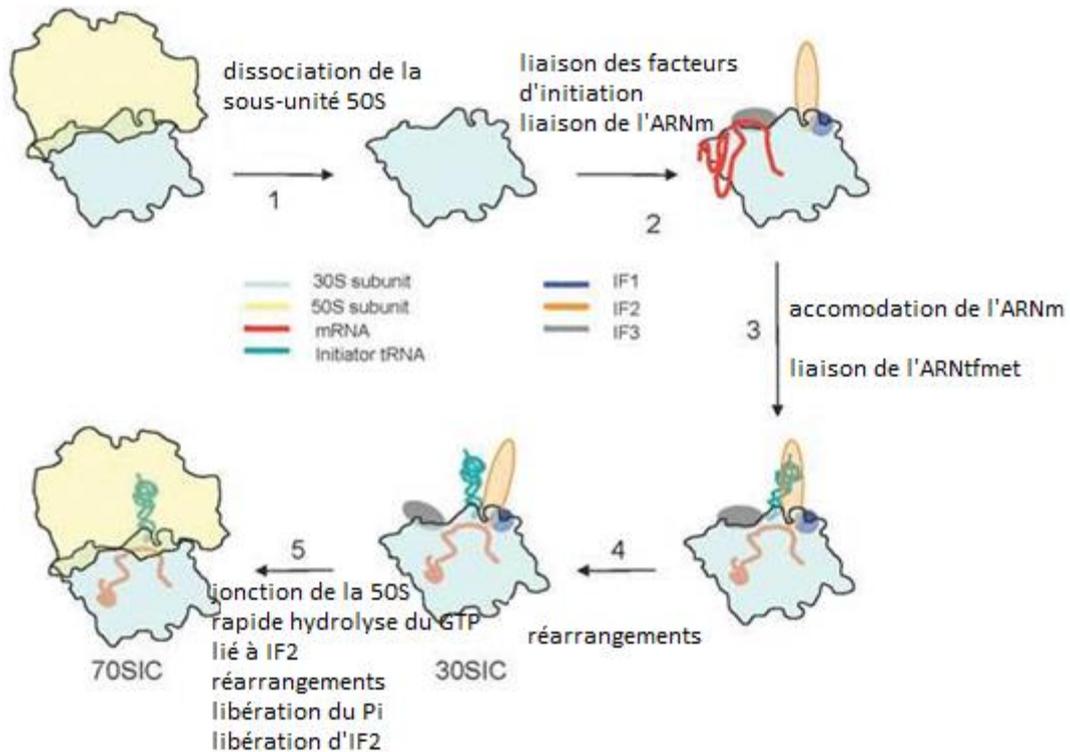


Figure 1.7 : Vue schématique du démarrage de la traduction chez les Bactéries  
 D'après Mechulam, Blanquet, et Schmitt 2011. Les partenaires impliqués sont schématisés en utilisant le code couleur au centre de la figure.

Etape initiale

Dans un premier temps, l'ARNm, les facteurs d'initiations IF1, IF3 (dont la structure et la fonction seront détaillées plus tard), IF2 ainsi que le fMet-ARNt<sup>Met</sup> sont assemblés sur la petite sous-unité 30S du ribosome, dans un ordre dépendant de la concentration en facteur dans le milieu (Tsai et al. 2012) afin de former le complexe de pré-initiation (30S-PIC).

Le facteur IF1 se fixe au niveau du site A de la 30S, empêchant ainsi le recrutement de l'ARNt initiateur au niveau du site A et par conséquent participe à sa fixation au site P (Dahlquist et Puglisi 2000; Carter et al. 2001).

IF3 stabilise également la liaison de l'ARNt initiateur au site P (C. Gualerzi, Risuleo, et Pon 1977) et maintient la dissociation des deux sous-unités du ribosome en se liant à la sous unité 30S. Ce positionnement de l'ARNt initiateur au niveau du site P de la 30S est réalisé selon un mode de liaison indépendant de son interaction avec le codon de l'ARNm (Tomsic et al. 2000). Le fMet-ARNt<sup>Met</sup> initiateur est spécifiquement reconnu par le facteur IF2 capable de se lier au ribosome.

## Le facteur IF2

Le facteur IF2 (97kDa chez *E. coli*) est une protéine G capable de lier un GTP et de l'hydrolyser en GDP et Pi via sa région centrale « G ». Malgré cela, la sélection du fMet-ARNt<sub>f</sub><sup>Met</sup> par IF2 est réalisée probablement directement sur le ribosome de façon indépendante de la présence de GTP (Sundari et al. 1976; Pon et al. 1985).

La zone d'interaction entre IF2 et le fMet-ARNt<sub>f</sub><sup>Met</sup> se ferait principalement via le groupe formyl du résidu méthionine et le domaine C-terminal de la protéine (Sundari et al. 1976; Li et al. 1996; Mayer et al. 2003) (voir figure 9 ci-dessous). La tige acceptrice peut aussi être un élément de reconnaissance par IF2 (Mayer et al. 2003) (voir figure 1.8).

Malgré tous ces contacts, le groupement formyl reste le déterminant majeur, comme nous avons pu le constater précédemment.

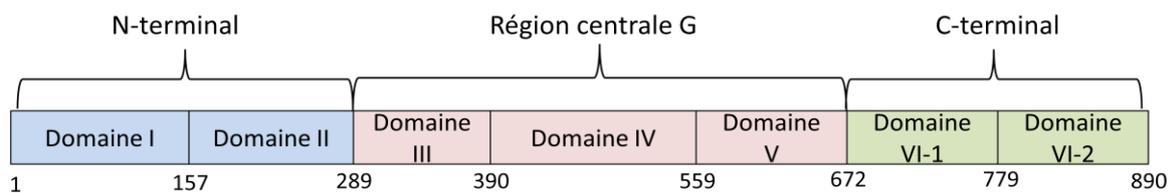


Figure 1.8 : Séquence primaire d'IF2. La numérotation est basée sur la séquence d'IF2 d'*E. coli*

De manière intéressante, ce facteur est homologue au facteur e/alf5B des eucaryotes et des archées (Moreno et al. 1999).

## Reconnaissance du codon de démarrage

Le ribosome est lié à l'ARNm par complémentarité de base entre une séquence nucléotidique riche en purine, appelée séquence Shine-Dalgarno (SD) (ou site de fixation du ribosome) se situant en amont du codon de démarrage sur l'ARNm et une séquence anti Shine-Dalgarno (aSD) située à l'extrémité 3' de l'ARN ribosomal (ARNr) 16S de la petite sous-unité 30S (Shine et Dalgarno 1974). La séquence SD canonique complète est UAAGGAGGU, mais elle est en générale trouvée sous forme partielle de 3 à 6 nucléotides consécutifs de cette séquence. L'appariement SD-aSD a pour conséquence de pré-positionner le complexe 30S-PIC au voisinage du codon de démarrage. Ce mécanisme d'appariement est une étape clé pour permettre au ribosome la reconnaissance du codon de démarrage AUG et le distinguer des codons AUG se trouvant à l'intérieur des cistrons qui codent pour des méthionines internes aux chaînes protéiques.

La reconnaissance par appariement du codon de démarrage AUG à l'anticodon du fMet-ARNt<sup>Met</sup> initiateur, préalablement recruté par IF2 sous sa forme lié au GTP, entraîne un changement de conformation du complexe de pré-initiation (30S-PIC) en un complexe d'initiation 30S (30S-IC) stable, fonctionnel et compétent pour le démarrage de la traduction. La validité de cet appariement serait vérifiée, en particulier grâce à IF3 qui aurait pour fonction de dissocier des complexes pour lequel le codon de l'ARNm et l'anticodon de l'ARNt seraient mésappariés (Hartz et al. 1990).

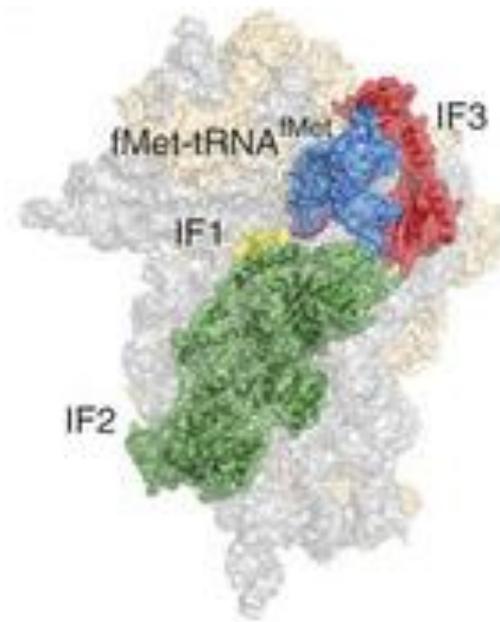


Figure 1.9 : Modèle structural du 30S-IC bactérien basé sur une reconstruction de cryo-EM. D'après Milón et al. 2012.

L'apport de différents modèles structuraux a permis de mieux établir les différents mécanismes de liaison des facteurs bactériens afin de former un complexe capable de démarrer correctement la traduction (voir figure 1.9)

#### Liaison au ribosome

Les structures mettant en jeu la petite sous-unité du ribosome ainsi que les différents facteurs de démarrage permettent de comprendre comment l'ARNt initiateur est dirigé préférentiellement vers le site P du ribosome.

En effet, IF1 se lie à la petite sous-unité 30S du ribosome au niveau du site A, bloquant ainsi son accès à l'ARNt (Carter et al. 2001). Le facteur IF3, venant se lier au niveau du site E du ribosome (voir figure 10 ci-dessus), pourrait avoir une fonction similaire à IF1 en bloquant l'accès du site E à

l'ARNt. L'ARNt initiateur n'aurait donc pas le choix et serait dirigé vers le site P du ribosome afin de pouvoir procéder à l'appariement codon-anticodon correct.

D'autre part, les facteurs pourraient agir de façon synergique les uns sur les autres. La présence d'IF1 sur la sous-unité 30S renforce la liaison d'IF2 sur le ribosome (Weiel et Hershey 1982). Ainsi, IF1 se lierait sur une région proche comprise entre le domaine III et VI d'IF2 (Moreno et al. 1999) et serait alors à proximité des domaines V et VI-2 du facteur IF2 tout comme de la tige anticodon de l'ARNt suggérant son rôle dans la sélection sur le ribosome de ce dernier.

Cependant, la structure de la petite sous-unité 30S avec IF1 et IF2 (Simonetti et al. 2008) montre qu'il n'existe aucune interaction directe entre ces deux facteurs, même si cette interaction serait rendue possible sur un complexe 70S.

#### Dissociation du complexe

Pour finir, le 30S-IC recrute la grande sous-unité 50S pendant que les facteurs IF1 et IF3 sont relâchés, formant ainsi le complexe 70S. Cette étape est catalysée par le facteur IF2. Elle nécessite l'hydrolyse du GTP en GDP et Pi. Lorsque le Pi provenant de l'hydrolyse est à son tour relâché, IF2 se dissocie du complexe (Antoun et al. 2003; C. O. Gualerzi et al. 1991). Ce complexe 70S stable et fonctionnel est alors prêt à traduire dans la bonne phase de lecture l'ARNm sélectionné.

### **1.2.2 Le démarrage eucaryote**

Tout comme chez les bactéries, le démarrage de la traduction chez les organismes eucaryotes est une étape limitante et essentielle de la survie cellulaire. Il nécessite un haut niveau de contrôle. Comme nous l'avons vu précédemment, la reconnaissance du bon codon de démarrage de la traduction nécessite la présence de 12 facteurs de démarrage chez les eucaryotes contre seulement 3 chez les bactéries.

Cette augmentation du nombre de facteurs peut s'expliquer par des différences majeures du mécanisme de démarrage entre les bactéries et les eucaryotes.

Au contraire des bactéries, les ARNm eucaryotes, sont monocistroniques et ne possèdent pas de séquence SD en amont du codon AUG. Les travaux de M. Kozak (M. Kozak 1981) ont amené à proposer un modèle de balayage de l'ARNm par le complexe ribosomal pour localiser le codon de démarrage.

Chez les eucaryotes, le complexe de démarrage est formé par l'assemblage de 12 facteurs protéiques de démarrage, d'un ARNt initiateur méthionylé (Met-ARNt<sub>i</sub><sup>Met</sup>) non formylé, et de la petite sous-unité du ribosome.

La reconnaissance du codon de démarrage n'est donc pas le fait d'un seul facteur, mais de plusieurs fonctionnant de façon coopérative, complexifiant ainsi la compréhension de ce mécanisme.

Malgré cette complexité, d'énormes avancées ont été réalisées dans la compréhension du rôle de chaque facteur dans la formation du complexe de démarrage au niveau du codon AUG (voir figure 1.10).

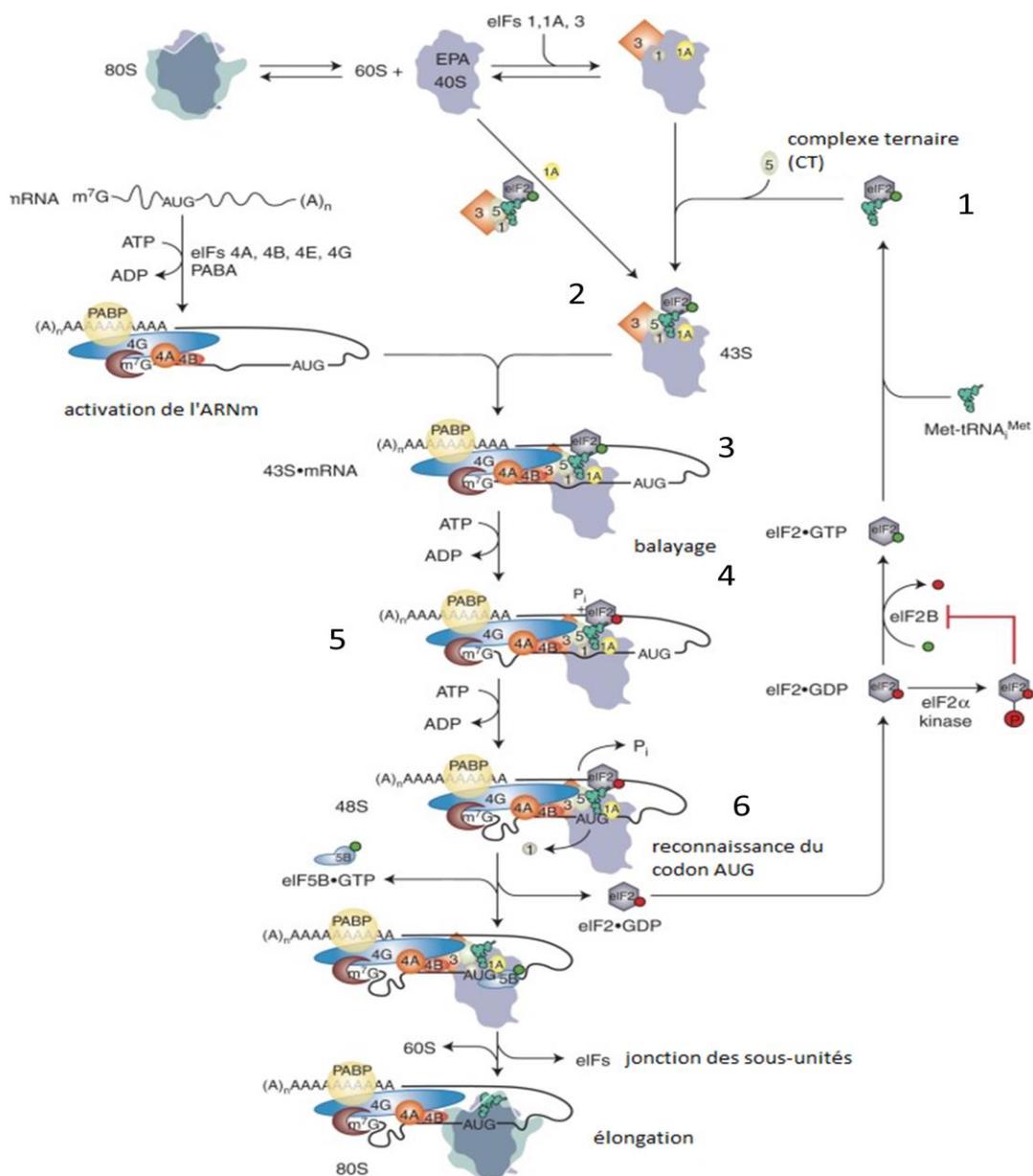


Figure 1.10 : Modèle du démarrage de la traduction Eucaryote. D'après Hinnebusch et Lorsch 2012

La succession des étapes du démarrage de la traduction peut être décrite de la manière suivante :

Chez les Eucaryotes, le Met-ARNt<sub>i</sub><sup>Met</sup> est apporté à la petite sous-unité du ribosome sous la forme d'un complexe ternaire (CT) formé de l'hétérotrimère eIF2 lié à une molécule de GTP et au Met-ARNt<sub>i</sub><sup>Met</sup> (voir étape 1). L'hétérotrimère eIF2 est organisé en sous-unité  $\alpha$ ,  $\beta$  et  $\gamma$ . 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 et Hannig 1996). La capacité de la sous-unité  $\gamma$  à lier le GTP a été définitivement montrée lors de la détermination de la structure cristallographique de la sous-unité  $\gamma$  (Schmitt et al. 2012).

Les facteurs eIF1, eIF1A, eIF5 et le complexe eIF3 accompagnent la liaison du CT à la sous-unité 40S, formant ainsi un complexe 43S de pré-démarrage (43S-PIC)(voir étape 2).

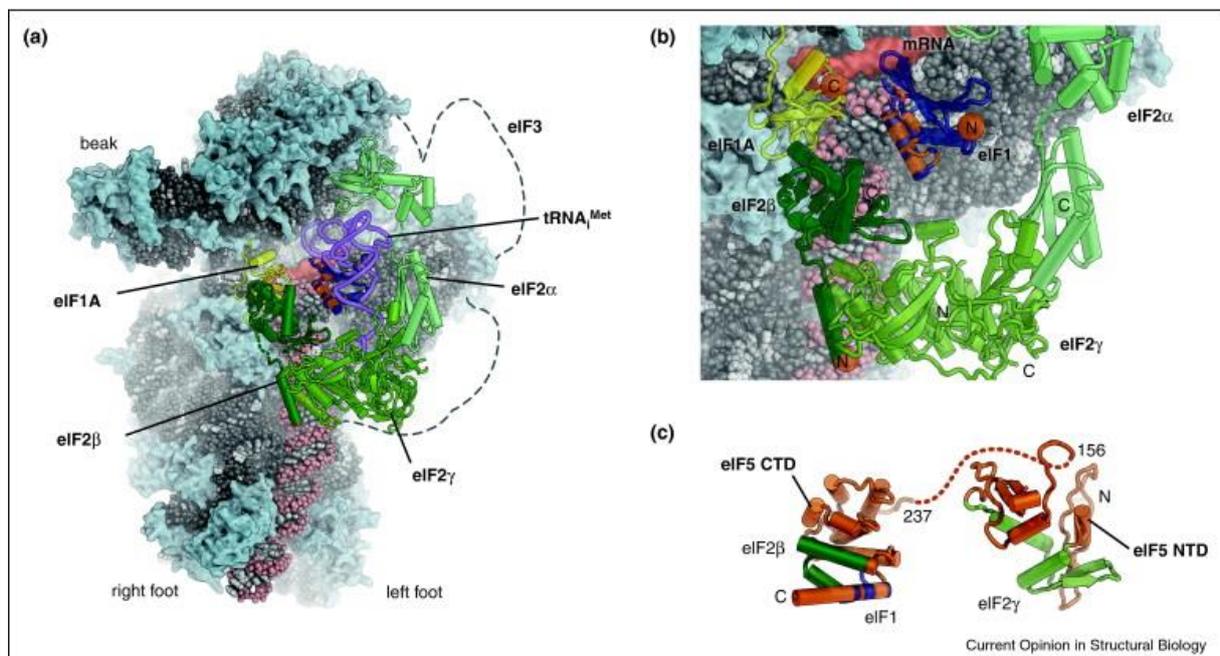


Figure 1.11 : Modèle de positionnement des facteurs de démarrage sur la petite sous-unité du ribosome dans le complexe de pré-initiation.

D'après Voigts-Hoffmann, Klinge, et Ban 2012

Le 43S-PIC se positionne alors au niveau de l'extrémité 5' maturée de l'ARNm. Cette étape est facilitée par le complexe eIF4F ainsi que par eIF4B qui se lie à l'extrémité 3' de l'ARNm (voir étape 3). Le complexe eIF4F est composé de la protéine de liaison à la coiffe eIF4E, liée à la protéine de liaison à l'ARN eIF4G via une hélicase à ARN, eIF4A et permet le désenroulement de la région structurée 5' non traduite (5'UTR) de l'ARNm. eIF4G, par le biais de son interaction avec eIF3, permettrait de faire

un pont entre l'ARNm circulaire et le complexe 43S-PIC, stimulant ainsi la liaison du complexe 43S à l'ARNm.

Lorsque l'appariement entre l'extrémité 5' de l'ARNm et le 43S-PIC est effectué, le complexe macromoléculaire balaye l'ARNm jusqu'au codon de démarrage (étape 4). Durant cette période balayage, le GTP lié à eIF2 est hydrolysé en GDP+Pi (voir étape 5). Cette hydrolyse est accélérée par eIF5, la « guanine activating protein ou GAP » (Algire, Maag, et Lorsch 2005).

Lorsqu'un codon AUG est trouvé dans un contexte nucléotidique correct, il y a appariement entre le codon AUG et l'anticodon CAU de l'ARNt<sub>i</sub><sup>Met</sup> (étape 6). Chez les mammifères, le contexte favorable au choix du codon est une séquence consensus nucléotidique mise en évidence comme modifiant l'efficacité du démarrage (M. Kozak 1986). La séquence consensus idéale est [5'-(A/G)NAUGG-3'] (Marilyn Kozak 2002). Cependant, cette complémentarité n'explique pas à elle seule l'arrêt du complexe sur ce codon AUG.

Après la reconnaissance du codon de démarrage, la petite sous-unité du ribosome change de conformation et eIF1 est dissocié de la 40S (Maag et al. 2005). Cette dissociation entrainerait la libération du Pi d'eIF2-GDP-Pi (Algire, Maag, et Lorsch 2005). Ainsi, l'hydrolyse du GTP serait rendue irréversible (Passmore et al. 2007). eIF2-GDP perdrait alors son affinité pour l'ARNt<sub>i</sub><sup>Met</sup> correctement accommodé et se séparerait du ribosome, laissant l'ARNt<sub>i</sub><sup>Met</sup> au site P de la sous-unité 40S.

Parallèlement, le complexe eIF2-GDP est recyclé grâce à l'échange du GDP en GTP par le facteur eIF2B.

Dans une dernière étape, une seconde GTPase, eIF5B (l'homologue du facteur bactérien IF2), facilite le recrutement de la grande sous-unité ribosomale, permettant ainsi la formation ainsi d'un complexe 80S, et le déplacement des facteurs de démarrage encore présents. L'hydrolyse du GTP lié à eIF5B est nécessaire pour permettre la dissociation d'eIF5B et eIF1A, libérant ainsi le site A du ribosome pour le premier ARNt élongateur.

### **1.2.3 Le démarrage archée**

Les analyses génomiques ont montré que le mécanisme de démarrage de la traduction des archées présentait à la fois des caractéristiques eucaryotes et des caractéristiques bactériennes.

Comme chez les bactéries, les ARN messagers peuvent être polycistroniques et ne portent pas de coiffe à l'extrémité 5' ni de queue poly-A à l'extrémité 3', mais ils possèdent souvent une séquence de type Shine-Dalgarno en amont du codon AUG de démarrage (Brown, Daniels, et Reeve 1989).

Cependant, ce modèle de démarrage de la traduction basé sur l'interaction SD-aSD n'est pas le seul possible chez les archées.

En effet, chez certaines archées, il existe des ARNm monocistroniques dépourvus de séquence 5'-UTR (région non traduite). La traduction de ces ARNm semblent, dans ce cas, nécessiter l'appariement direct entre le codon de démarrage et l'anticodon porté par l'ARNt<sup>Met</sup> (Benelli, Maone, et Londei 2003).

Le premier facteur de démarrage putatif trouvé chez les archées est homologue à un facteur eucaryote (eIF5B) (Bartig et al. 1992), laissant apparaître des similarités entre le processus de démarrage de la traduction archée et eucaryote.

Comme chez les eucaryotes, l'ARNt initiateur méthionylé n'est pas formylé et il est conduit au ribosome par un homologue de la protéine hétérotrimérique eucaryote eIF2, notée aIF2.

Ce dernier facteur, aIF2, est également formé de trois sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$ . Tout comme son orthologue eucaryote, il permet, sous sa forme liée au GTP, d'apporter l'ARNt méthionylé à la sous-unité 30S du ribosome, formant ainsi un complexe ternaire aIF2:GTP:ARNt<sup>Met</sup>.

Les homologues des deux petites protéines eIF1 et eIF1A essentielles à la formation du complexe de démarrage eucaryote sont également présentes chez les archées et sont nommées aIF1 et aIF1A. Le facteur aIF1 (que nous détaillerons dans une prochaine partie) présente une forte homologie au facteur eIF1, ce qui laisse présager une fonction similaire au facteur eucaryote dans le démarrage de la traduction. Il en est de même pour la fonction du facteur aIF1A, orthologue au facteur eucaryote eIF1A.

D'autres facteurs comme eIF3, eIF5, eIF4 participant à la formation du complexe de démarrage eucaryote sont absents chez les archées.

Pour toutes ces raisons, le démarrage de la traduction chez les archées apparaît comme une version simplifiée du démarrage de la traduction des eucaryotes. En effet, dans les deux domaines du vivant, un cœur structural similaire formé des facteurs e/aIF1, e/aIF1A et e/aIF2:GTP:ARNt intervient dans le démarrage (figure 1.12 ci-dessous).

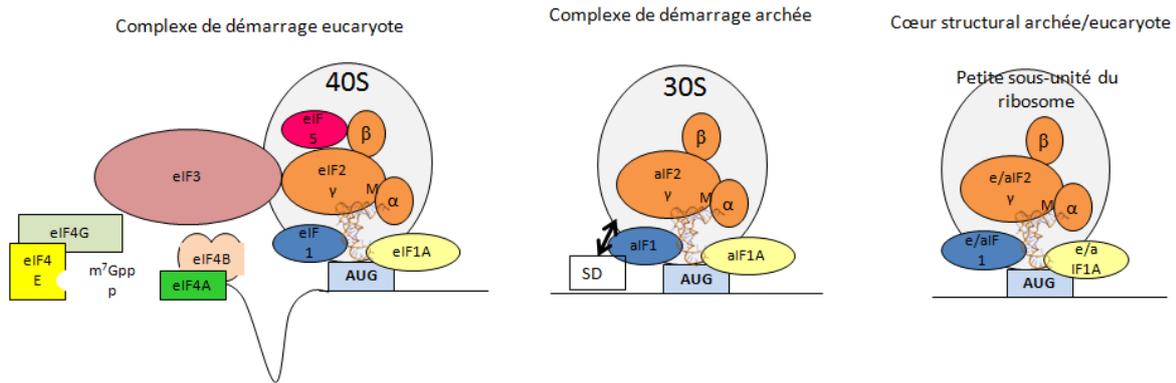


Figure 1.12 : Schéma des complexes de démarrage eucaryote (gauche) et archéen (au centre). Le cœur structural est représenté à droite.

Ainsi, malgré les différences constatées lors du démarrage de la traduction et l'implication de différents acteurs selon le domaine du vivant observé, un groupe de 3 protéines homologues, se distingue chez les eucaryotes et les archées comme étant essentiel à la formation d'un complexe de démarrage fonctionnel (voir tableau ci-dessous).

De façon intéressante, deux homologues de ces trois protéines sont aussi retrouvés chez les bactéries, IF1 homologue d'e/aIF1A et le domaine C-terminal d'IF3 homologue à e/aIF1, soulignant leur caractère universel et fondamental dans le démarrage de la traduction.

Les facteurs du complexe cœur du démarrage de la traduction	
archées	eucaryotes
aIF2 (95 kDa)	eIF2 (127.3 kDa)
aIF1A (13 kDa)	eIF1A (16.4 kDa)
aIF1 (11.4 kDa)	eIF1 (12.7 kDa)

Les poids moléculaires correspondent respectivement à ceux des facteurs de *Pyrococcus abyssi* et de l'*Homo sapiens sapiens*.

## 1.3 Le complexe cœur du démarrage de la traduction

### 1.3.1 Le facteur hétérotrimérique e/1F2

#### 1.3.1.1 Identification du rôle d'e1F2 dans la sélection du codon de démarrage

A l'origine de la découverte des acteurs protéiques impliqués dans le démarrage de la traduction eucaryote, 2 facteurs ont été caractérisés comme jouant un rôle dans cette étape de façon indépendante de la présence d'un ARNm et permettant la liaison d'un ARNt<sub>i</sub><sup>Met</sup> sur la petite sous-unité du ribosome (Schreier et Staehelin 1973).

Le premier acteur est le facteur multimérique e1F3, spécifique du démarrage de la traduction eucaryote et ne bénéficiant pas d'homologues chez les archées.

Le second acteur protéique est le facteur e1F2, déjà caractérisé comme pouvant former un complexe ternaire avec une molécule de GTP et un ARNt<sub>i</sub><sup>Met</sup> (Levin, Kyner, et Acs 1971; Dettman et Stanley 1972; Levin, Kyner, et Acs 1973)

Par la suite, le rôle des facteurs protéiques dans la spécificité du démarrage de la traduction a été mis en évidence grâce à la caractérisation de mutants de levures devenues capables de démarrer la traduction d'un ARNm portant un codon de démarrage muté. Le crible génétique utilisé consistait à sélectionner des mutants supprimeurs d'une mutation du codon de démarrage du gène codant pour la protéine His4, présentant un codon AUU à la place d'un codon AUG.

Ces mutations « supprimeurs » se retrouvent dans 5 gènes (*sui1* à *sui5*), dont 3 s'avèrent coder pour les sous-unités du facteur e1F2.

Les gènes *sui2*, *sui3* et *sui4* (ou *gcd11*) codent respectivement pour les sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$  du facteur e1F2 (Cigan et al. 1989; Donahue et al. 1988; Huang et al. 1997).

Ainsi, en plus de leur rôle dans la liaison de l'ARNt initiateur, les trois sous-unités d'e1F2 ont également une fonction dans la sélection du codon de démarrage correct. En effet, des mutations introduites volontairement cette fois dans une des trois sous-unités de ce facteur peuvent permettre le démarrage de la traduction sur un mauvais codon de démarrage (Williams, Hinnebusch, et Donahue 1989; Castilho-Valavicius, Yoon, et Donahue 1990; Dorris, Erickson, et Hannig 1995).

Chez les eucaryotes, l'hydrolyse du GTP en GDP par e1F2 est assistée par la protéine GAP activatrice, e1F5. Chez les archées cependant, aucune protéine homologue à e1F5 n'a pu être identifiée, laissant supposer que l'activité GTPase d'e1F2 ne nécessiterait aucune assistance. De même, l'échange

GDP/GTP assisté par eIF2B chez les eucaryotes semble être spontané chez les Archées. De manière intéressante, des mutants Sui- ont été obtenus dans le gène codant pour eIF5 (*sui5*). Ainsi, l'activité d'hydrolyse du GTP sur eIF2 paraît également couplée à la reconnaissance du codon de démarrage.

### 1.3.1.2 Le complexe ternaire (CT) e/aIF2-GTP-ARN<sup>i</sup>Met

La récente détermination à 5 Å de résolution de la structure cristallographique du complexe aIF2 de l'archée *Sulfolobus solfataricus* lié à l'analogue non hydrolysable du GTP, le GDPNP ainsi qu'à l'ARNt initiateur d'*E. coli* permet de décrire la liaison de l'ARNt au facteur aIF2 (Schmitt et al. 2012).

Dans l'hétérotrimère, la sous-unité  $\gamma$  forme le cœur du complexe. Elle interagit avec les sous-unités  $\alpha$  et  $\beta$ , tandis que celles-ci n'interagissent pas entre elles (Schmitt, Naveau, et Mechulam 2010) (voir figure 1.13)

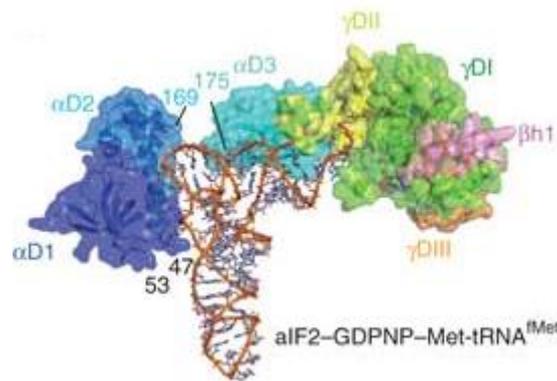


Figure 1.13 : modèle du complexe ternaire de démarrage de la traduction formé par le facteur aIF2 de *Sulfolobus solfataricus* et l'ARNt<sup>Met</sup> d'*E. coli*. D'après Schmitt et al. 2012

On peut remarquer grâce à ce modèle que l'hélice acceptrice de l'ARNt<sup>Met</sup> est enveloppée par les sous-unités  $\alpha$  et  $\gamma$  d'aIF2. Le coude de l'ARNt est lié par les domaines 1 et 2 de la sous-unité  $\alpha$  (en bleu dans la figure). Le petit sillon de la tige acceptrice de l'ARNt est conduit le long du feuillet  $\beta$  du domaine 3 d' $\alpha$  et du domaine II de  $\gamma$  d'une part, et par l'interface des domaines I et II de  $\gamma$  d'autre part. Pour finir, l'extrémité de la tige acceptrice de l'ARNt, portant la méthionine estérifiée sur la base A<sub>76</sub>, est liée dans une cavité du domaine II de la sous-unité  $\gamma$ .

De manière intéressante, il faut souligner qu'en dépit d'une très grande similarité entre la structure de la sous-unité  $\gamma$  et celle du facteur d'allongement de la synthèse peptidique EF1-A, le mode de liaison de l'ARN de transfert par ces deux protéines est complètement différent. Ceci suggère que la manière dont le complexe ternaire de démarrage, e/aIF2:ARNt:GDPNP, approche le ribosome est, elle aussi, différente de celle utilisée par le facteur d'allongement. Finalement, un modèle de liaison

du complexe ternaire  $\alpha$ IF2-GTP-ARNt au ribosome a été directement déduit de cette structure (voir figure 1.14)

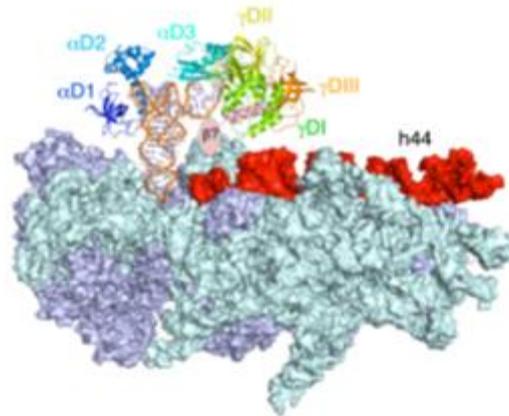


Figure 1.14 : modèle de liaison du complexe ternaire eIF2:ARNt:GTP sur la petite sous-unité du ribosome 40S

### 1.3.2 Les facteurs IF1 et $\alpha$ /eIF1A

Comme nous avons pu le voir précédemment, les facteurs IF1 bactériens et  $\alpha$ /eIF1A archée et eucaryotes sont des facteurs orthologues.

#### 1.3.2.1 Le facteur IF1

Le facteur IF1 bactérien est un petit facteur de liaison à l'ARN (71 acides aminés chez *E. coli*) codé par le gène *infA* (Sands et al. 1987) et essentiel à la survie cellulaire d'*E. coli* (Cummings et Hershey 1994), suggérant qu'au moins une de ses fonctions est cruciale in vivo. Sa structure, tout d'abord déterminée par RMN (Sette et al. 1997), puis par cristallographie, est caractérisée par 5 brins organisés en tonneau  $\beta$  avec une grande boucle flexible reliant les brins 3 et 4. Cette structure est caractéristique d'un motif oligomère de liaison aux acides nucléiques ou OB-fold (Sette et al. 1997) de la famille des protéines S1 (Bycroft et al. 1997) (voir figure 1.15).

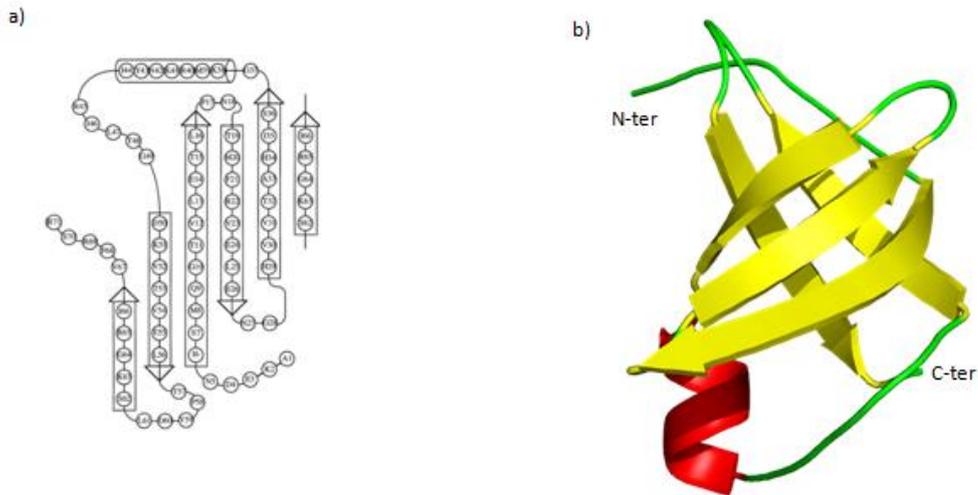


Figure 1.15 : (a) Topologie du repliement d'IF1 d'après Sette et al. 1997 M. Sette et al., 1997 (b) Structure 3D d'IF1 d'après Carter et al. 2001

La détermination de cette structure a permis de suggérer que cette protéine avait un rôle dans le démarrage de la synthèse protéique, de par ses interactions possibles avec les acides nucléiques. Par la suite, l'utilisation de sondes chimiques sur l'ARN ribosomal et l'analyse de mutants de ribosome ont mis en évidence que les bases  $C_{1407}$ ,  $A_{1408}$ ,  $A_{1492}$  et  $A_{1493}$  de l'ARNr16S jouaient un rôle dans la liaison d'IF1 à la petite sous-unité du ribosome (Dahlquist et Puglisi 2000), et de proposer la liaison de ce facteur au site A de la petite sous-unité.

En 2001, la résolution de la structure cristallographique du complexe IF1-30S de *T. thermophilus* (Carter et al. 2001) a permis de comprendre plus finement les interactions mises en jeu lors de la liaison d'IF1 au niveau du site A du ribosome, et a fourni une vue des changements conformationnels se produisant au sein de la petite sous-unité ribosomique lors de la liaison de ce facteur (voir figure 1.16)

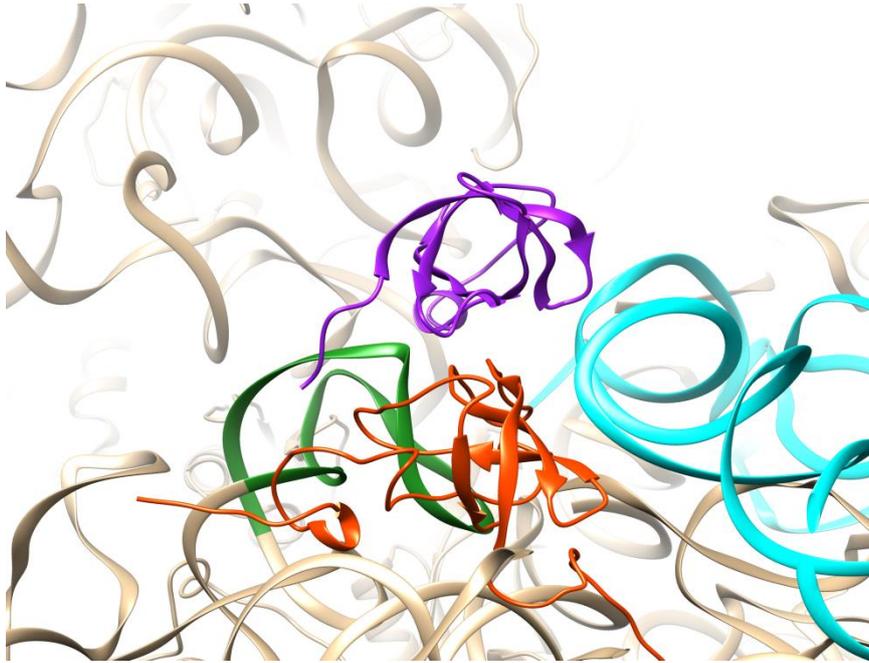


Figure 1.16 : Vue structurale de l'interaction entre IF1 et la sous-unité 30S de *Thermus thermophilus*. IF1 est en violet, l'hélice 44 en cyan, la boucle 530 en vert et la protéine S12 en orange. D'après Carter et al.

2001

La liaison d'IF1 produit un changement conformationnel de la sous-unité 30S qui affecte l'exposition à des réactifs chimiques de bases distantes telles que A<sub>908</sub> et A<sub>909</sub> (Dahlquist et Puglisi 2000). Ces expositions reflètent de petits mouvements de rotation de la tête, la plateforme et de l'épaule de la 30S vers le site A. Ces mouvements sont visibles si l'on compare la structure du complexe 30S-IF1 avec celle de la sous-unité 30S seule dans les mêmes conditions (Wimberly et al. 2000).

IF1 se lie à la sous-unité 30S dans une cavité située entre l'hélice H44, la boucle G<sub>530</sub> et la protéine S12 (voir figure 17 ci-dessus). La surface d'IF1 interagissant avec la sous-unité 30S est riche en résidus basiques alors que les résidus acides sont plutôt en interaction avec le solvant.

La liaison d'IF1 induit un retournement des bases A<sub>1492</sub> et A<sub>1493</sub> de l'hélice H44 dans une poche d'IF1 définie par la boucle 17-25. Cet internement des bases dans IF1 explique pourquoi celles-ci ne sont pas sujettes à des modifications chimiques (Moazed et al. 1995).

Des analyses par mutagenèse dirigée ont de plus montré l'importance de l'arginine 69 située au niveau de l'extrémité C-terminale d'IF1 pour la liaison du facteur à la sous-unité 30S (C. O. Gualerzi et al. 1989).

La liaison d'IF1 au niveau du site A du ribosome préviendrait la liaison prématurée d'un ARNt à ce niveau au moment du démarrage de la traduction. D'autre part, la présence du facteur IF1 seul n'affecte pas l'efficacité de la liaison de l'ARNt initiateur au site P. La présence d'IF2 ou d'IF3 est nécessaire pour détecter un effet d'IF1 dans la formation du complexe de pré-initiation 30S

(Wintermeyer et Gualerzi 1983). En effet, le facteur IF1 pourrait permettre l'augmentation de l'effet déstabilisateur d'IF3 sur l'ARNt<sub>i</sub><sup>Met</sup>, tout en augmentant l'effet stabilisateur d'IF2 en absence d'IF3. IF1 prévient donc la dissociation d'IF2 de la 30S de façon indépendante de la présence d'IF3 (Stringer, Sarkar, et Maitra 1977). De plus, au moment de l'étape d'assemblage avec la grande sous-unité, il contribuerait à pouvoir promouvoir l'hydrolyse du GTP sur IF2. Pour finir, des expériences de marquage d'affinité (cross-linking) laissent penser qu'une interactions directe pourrait exister entre IF1 et IF3 d'une part et entre IF1 et IF2 d'autre part (Boileau et al. 1983), bien que les structures disponibles d'IF3 sur le ribosome n'aillent pas dans ce sens (voir partie IF3).

### 1.3.2.2 Structures des facteurs a/eIF1A

#### 1.3.2.2.1 Le facteur eucaryote eIF1A :

Le facteur eIF1A fait partie des premiers facteurs eucaryotes reconnus comme participant au démarrage de la traduction.

Historiquement, différents test d'assemblage de complexes ribosomiaux contenant un ARNm, un Met-ARNt<sub>i</sub><sup>Met</sup>, les 7 facteurs de démarrage purifiés ainsi que de l'ATP et du GTP avaient été entrepris. L'omission d'un des composants à chaque essai avait permis de mettre en avant le rôle de chacun de ces facteurs. Ainsi, l'importance du facteur eIF1A (anciennement appelée eIF4C) dans la liaison de l'ARNm à la 40S a été mise en lumière (Trachsel et al. 1977). Les résultats suggèreraient une interaction possible entre le facteur eIF1A et la petite sous-unité du ribosome.

Par la suite, des purifications de complexes ribosomiaux sur colonne Sépharose 6B (colonne de chromatographie par tamis moléculaire) puis sur gradient de sucrose, avaient permis de révéler que le facteur eIF1A était capable de se lier à la petite sous-unité 40S du ribosome, de façon quasi stoechiométrique, lorsque les facteurs eIF1, eIF2, eIF3 ainsi qu'un ARNt<sub>i</sub><sup>Met</sup> et un ARNm étaient présents dans le milieu (Thomas et al. 1980).

Finalement, l'identification d'eIF1A comme étant un facteur essentiel du démarrage de la traduction est tout à fait cohérente avec le fait qu'il soit retrouvé comme l'un des facteurs eucaryotes les plus conservés (>65%) (Dever et al. 1994) et nécessaire à la viabilité de la cellule (Wei, MacMillan, et Hershey 1995).

Le facteur eIF1A est, comme son orthologue bactérien IF1, un petit facteur possédant un site de reconnaissance des acides nucléiques (OB-fold) central. Cependant, eIF1A possède en plus un domaine hélicoïdal C-terminal et des extensions non structurées en N-terminal et C-terminal (Battiste et al. 2000).

Dans l'ensemble, ce facteur est très chargé avec environ 46% de ces acides aminés qui s'avèrent être des lysines, arginines, aspartates ou encore des glutamates. Les séquences bactériennes et eucaryotes de ce facteur sont identiques à environ 21% (Kyrpides et Woese 1998). Ce pourcentage de similarité laisse sous-entendre des fonctions similaires des facteurs IF1 et eIF1A dans la cellule.

### 1.3.2.2.2 Le facteur archée aIF1A :

Les premiers séquençages de génomes d'archées (par exemple le génome de l'archée *M. jannaschii*, (Bult et al. 1996)) ont permis d'envisager une stratégie de recherche de facteurs de démarrage de la traduction homologues aux facteurs eucaryotes en utilisant des algorithmes de types BlastP (Basic Local Alignment Search Tool) (Altschul et al. 1990). Grâce à cela, différents facteurs, dont le facteur aIF1A sont découverts chez les archées. En effet, la séquence du facteur archée possède 38% d'identité avec celui du facteur eucaryote (et 30% avec le facteur IF1 bactérien) (Kyrpides et Woese 1998)(voir figure 1.17)

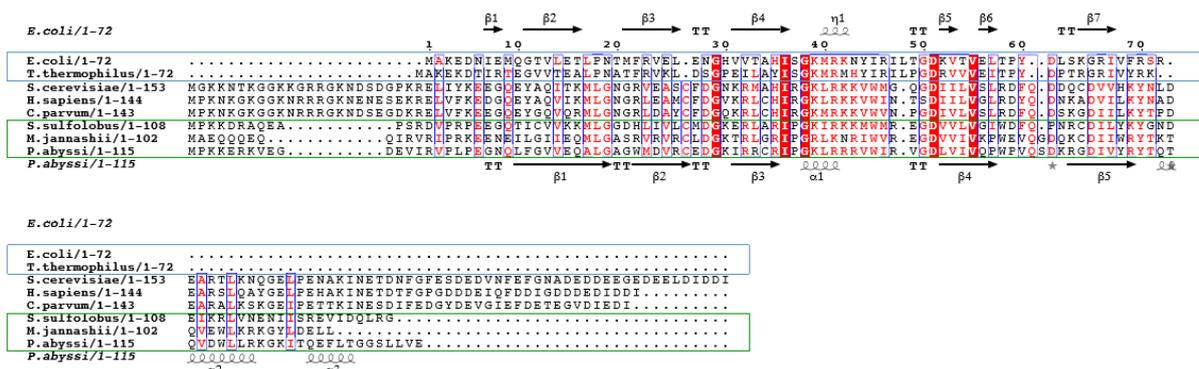


Figure 1.17 : Alignement des séquences des facteurs IF1, aIF1A et eIF1A

Le facteur aIF1A possède aussi une extension N-terminale par rapport à son homologue bactérien IF1. Celle-ci est toutefois plus courte que chez les eucaryotes. D'autre part, la partie C-terminale acide est aussi plus courte chez le facteur archée que chez le facteur eucaryote. (voir figure 1.18). En

effet, on trouve  $\alpha 2$  et  $\alpha 3$  comme chez eIF1A, par contre l'extrémité C-terminale acide eucaryote est absente ou très courte.

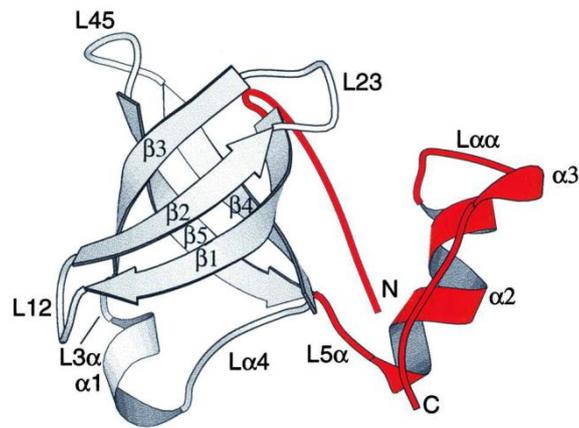


Figure 1.18 : Diagramme d'eIF1A d'après Battiste et al. 2000. Résidus 25 à 117. Le domaine OB et l'hélice alpha sont colorés en gris et rouge respectivement.

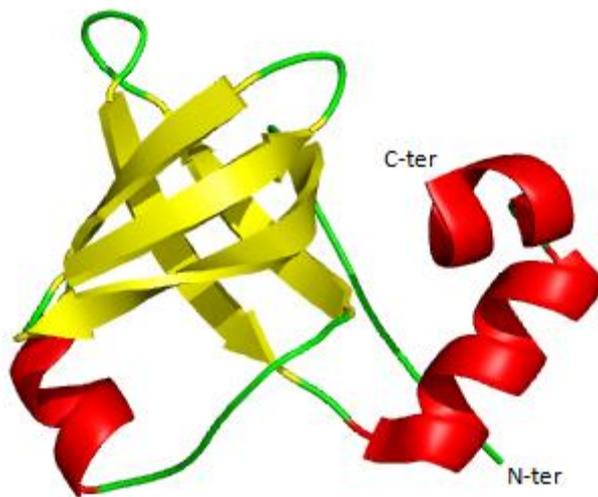


Figure 1.19 : Structure cristallographique d'aIF1A de *Pyrococcus abyssi*

Ces deux facteurs archée et eucaryote pourraient donc avoir quelques fonctions divergentes.

### 1.3.3 Structure des facteurs IF3 et a/eIF1

#### 1.3.3.1 Le facteur bactérien IF3

Le facteur IF3 est un petit monomère (180 acides aminés chez *E. coli*) basique essentiel à la survie cellulaire (Olsson et al. 1996), codé par le gène *InfC* (Sacerdot et al. 1982).

Ce facteur peut être divisé en deux domaines structurés relativement compacts de tailles presque égales, séparés par un peptide de liaison (Kycia et al. 1995). Ces deux domaines sont reliés par une séquence riche en lysine très conservée chez les bactéries et n'entrent pas en contact l'un avec l'autre, même lorsqu'IF3 est lié à la sous-unité 30S (Moreau et al. 1997; de Cock, Springer, et Dardel 1999).

Le domaine C-terminal est composé de 4 brins beta soutenus par deux hélices alpha (Garcia et al. 1995b) et le domaine N-terminal de 4 brins beta antiparallèles entourés par deux hélices alpha parallèles (Garcia et al. 1995a). Les deux domaines d'IF3 peuvent interagir avec le ribosome (voir figure 1.20).

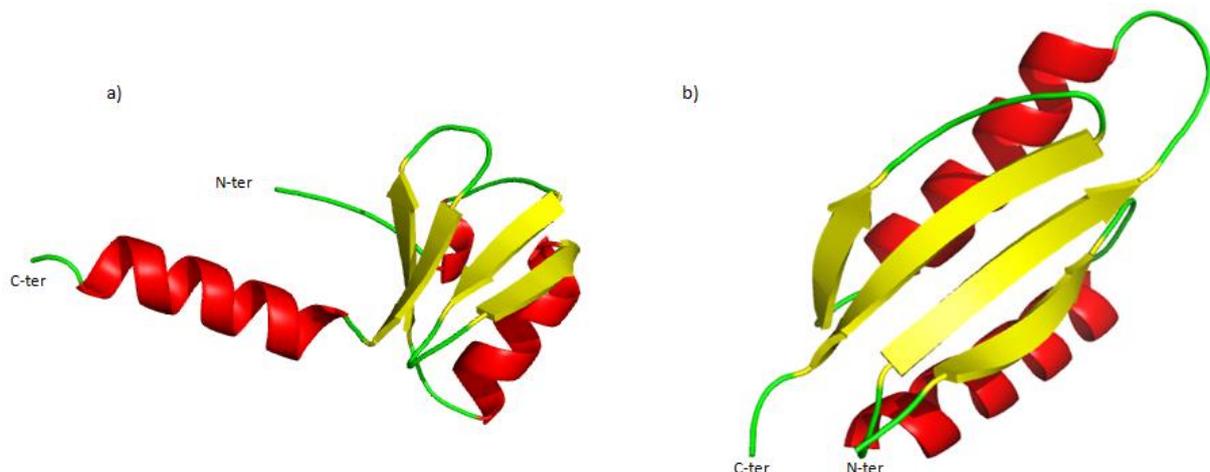


Figure 1.20 : Structures cristallographiques des domaines N (a) et C (b) d'IF3 de *G. stearothermophilus*. D'après Biou, Shu, et Ramakrishnan 1995

De manière intéressante, toutes les fonctions d'IF3 peuvent être montrées *in vitro* en utilisant uniquement le domaine C-terminal isolé. Il est proposé que le domaine N-terminal serve uniquement à la stabilisation de l'interaction du domaine C d'IF3 et de la petite sous-unité 30S (Petrelli et al. 2001).

Les résidus Y<sub>107</sub> et K<sub>110</sub>, ainsi qu'un groupe de résidus arginines situés au niveau du domaine C-terminal d'IF3 d'*E.coli* sont impliqués dans sa liaison au ribosome (De Bellis et al. 1992). De manière intéressante, seul le domaine C-terminal d'IF3 est homologue au facteur eIF1 eucaryote.

La résolution par Cryo-EM d'une première structure du complexe IF3-30S de *T. thermophilus* à 27Å de résolution propose de situer le facteur IF3 entre la plateforme et la tête de la petite sous-unité du ribosome (McCutcheon et al. 1999). Cependant, l'alignement de cette structure avec les récentes structures à hautes résolutions de la sous-unité 30S isolée place la densité attribuée à IF3 au niveau des protéines ribosomales, rendant ainsi le placement proposé peu vraisemblable. Plus récemment, la structure cristallographique du complexe d'IF3-30S (Pioletti et al. 2001) situerait le domaine C-terminal d'IF3 sur la plateforme du ribosome. Cependant, si le domaine C-terminal d'IF3 est modélisé sur cette zone dans une structure de sous-unité 30S contenant un ARNm, on remarque que l'hélice  $\alpha$  d'IF3 se trouverait au niveau de l'extrémité 3' de l'ARNr 16S, et empêcherait ainsi l'appariement SD/aSD. Ce placement serait donc délétère au démarrage de la traduction. Finalement, cette structure ne permet pas de proposer un placement du domaine N-terminal.

En 2005, la structure par Cryo-EM de la sous-unité 30S du ribosome en complexe avec les facteurs de démarrage de la traduction bactérien (Allen et al. 2005) propose de placer le domaine C-terminal d'IF3 entre la plateforme et la tête de la sous-unité. Une fois encore, la résolution ne permet pas de placer le domaine N-terminal, et la densité électronique associée au domaine C-terminal reste peu convaincante.

En conclusion, il apparaît que ces structures ne permettent pas de déterminer avec certitude la zone d'interaction d'IF3 avec la petite sous-unité 30S du ribosome.

Les expériences les plus convaincantes pour caractériser le site de liaison d'IF3 à la petite sous-unité du ribosome ont en fait été obtenues grâce à des expériences de protection de l'ARNr 16S en présence d'IF3 (Moazed et al. 1995) et d'empreinte chimique utilisant les radicaux hydroxyles (Dallas et Noller 2001). Ainsi, il est proposé que le domaine C terminal d'IF3 interagisse avec les bases des hélices 23, 24 et 45 de l'ARNr 16S. Pour finir, la liaison d'IF3 à la 30S pourrait contribuer à faire bouger les bases G<sub>1338</sub> et A<sub>1339</sub>, en contact direct avec le petit sillon de la tige anticodon de l'ARNt<sub>f</sub><sup>Met</sup> initiateur.

### 1.3.3.2 Structure des facteurs eucaryotes et archées e/aIF1

#### 1.3.3.2.1 Le facteur eucaryote eIF1

Tout comme le facteur eIF1A, le facteur eIF1 a pu être purifié et caractérisé comme étant un facteur de démarrage de la traduction dès 1977 grâce à des purifications de fractions ribosomales capables de démarrer la traduction d'un ARNm codant pour la globine de lapin (Schreier, Erni, et Staehelin 1977).

Son implication dans la sélection du bon codon de démarrage de la traduction a été mise en évidence, tout comme pour le facteur eIF2, par la découverte des mutants de levure capables de restaurer l'expression du gène codant pour HIS4 possédant un codon de démarrage muté (Yoon et Donahue 1992).

En effet, parmi les 5 gènes dans lesquels se retrouvent les mutations (vu précédemment), le gène SUI1 s'avère coder pour une protéine essentielle du démarrage de la traduction fonctionnant de concert avec le facteur eIF2. Il a ensuite été montré que le gène SUI1 codait pour le facteur eIF1 (Kasperaitis, Voorma, et Thomas 1995; Thomas et al. 1980).

eIF1 est une protéine de 113 résidus chez l'humain très conservée dans le domaine eucaryote. On constate 59% d'identité entre la séquence d'eIF1 humain et celle d'eIF1 de levure (Kyripides et Woese 1998).

La résolution de la structure d'eIF1 par RMN a montré un facteur monomérique globulaire composé de deux hélices  $\alpha$  d'un côté et de 5 brins  $\beta$  de l'autre (Fletcher et al. 1999)(voir figure 1.21).

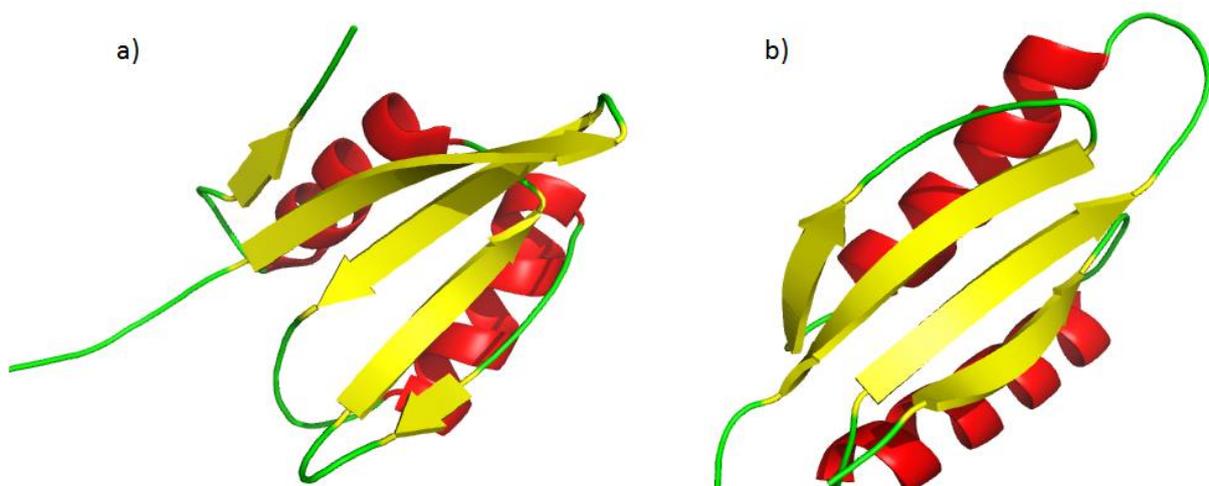


Figure 1.21 : (a) Structure 3D du facteur eIF1 humain d'après Lomakin et Steitz 2013 et (b) du domaine C-terminal d'IF3 d'après Biou, Shu, et Ramakrishnan 1995

Cette structure est homologue au domaine C-terminal du facteur bactérien IF3.

Des expériences d'empreintes utilisant des radicaux hydroxyyles (hydroxyl radical footprinting) ont permis de montrer que le facteur eIF1 était capable de se lier au ribosome entre la plateforme de la petite sous-unité et l'ARNt<sup>Met</sup> (Lomakin et al. 2003). Le domaine C-terminal se situerait au niveau de la zone de liaison de l'hélice 69 de l'ARNr 23S, expliquant ainsi son rôle dans le blocage de l'association des sous-unités.

De manière intéressante, la position identifiée pour eIF1 sur la 40S est similaire à celle mise en évidence pour le domaine C-terminal d'IF3 des bactéries (Dallas et Noller 2001; Lomakin et al. 2003). En 2011, la détermination de la structure cristallographique d'eIF1 en complexe avec la petite sous-unité 40S du ribosome (Rabl et al. 2011) a permis de révéler que ce facteur se liait à la 40S sur un site au sommet des hélices 24 et 44 de l'ARNr 18S du ribosome sur le côté droit de la plateforme à proximité du site P *via* une région chargée positivement. La position de liaison d'eIF1 à la 40S « amène » une boucle riche en résidus basiques (appelée boucle basique) à proximité du canal de l'ARNm.

### 1.3.3.2.2 Le facteur archée aIF1

Des alignements de séquences ont permis de retrouver un homologue au facteur de démarrage eucaryote eIF1 chez les archées : le facteur aIF1. Ces séquences ont un pourcentage d'identité élevé d'environ 30% (Kyrpides et Woese 1998). Cependant, le facteur archée ne possède pas à son extrémité C-terminale la séquence « HGF » très conservée chez le facteur eucaryote (voir figure 1.22).

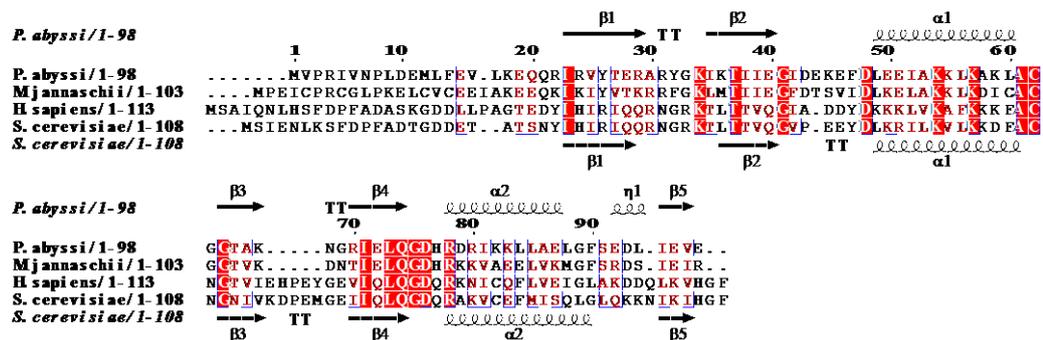


Figure 1.22 : Alignement des séquences des facteurs de démarrage aIF1 archées et eIF1 eucaryote.

### **1.3.4 Fonction des facteurs eucaryotes eIF1 et eIF1A**

L'une des premières étapes du démarrage de la traduction est la liaison des facteurs eIF1 et eIF1A à la petite sous-unité du ribosome. L'action synergique de ces deux facteurs lors du démarrage de la traduction a pu être montrée par des expériences d'empreinte du ribosome ou toeprinting avec un complexe de démarrage assemblé in vitro contenant la petite sous-unité du ribosome et différents facteurs (eIF2, eIF3, eIF4A, eIF4B et eIF4F) ainsi que l'ARNm naturellement coiffé de la  $\beta$ -globine et l'ARNt<sup>Met</sup> (Pestova, Borukhov, et Hellen 1998). De plus, des expériences d'anisotropie de fluorescence montrent que la liaison de ces facteurs à la petite sous-unité du ribosome se fait de façon indépendante à la présence des autres facteurs et est coopérative (Maag et Lorsch 2003).

La découverte de mutations entraînant des phénotypes Sui- dans les facteurs eIF1 et eIF1A ainsi que des expériences de toeprinting ont permis de mettre en évidence leur implication dans la sélection du bon codon de démarrage AUG. En effet, eIF1 est essentiel pour empêcher le démarrage de la traduction à partir d'un codon non AUG ou un codon AUG dans un mauvais contexte (Yoon et Donahue 1992; Pestova et Kolupaeva 2002; Valásek et al. 2004).

#### **1.3.4.1 Rôle des facteurs eIF1 et eIF1A dans le balayage de l'ARNm par le complexe de démarrage de la traduction**

Chez les eucaryotes, le complexe de démarrage de la traduction se lie au niveau de la coiffe guanosine en 5' sur l'ARNm. La découverte du codon de démarrage AUG sur l'ARNm est rendue possible grâce au balayage de celui-ci par le complexe ribosomal sur lequel sont fixés les différents acteurs du démarrage de la traduction. Des expériences de toeprinting ont permis de mettre en évidence le rôle essentiel des facteurs eIF1 et eIF1A dans le balayage de l'ARNm (Pestova, Borukhov, et Hellen 1998). En effet, en leur absence, un complexe ribosomal 43S de pré-initiation incubé en présence d'ATP et des facteurs de démarrage eIF4A, eIF4B et eIF4F reste lié sur une région non traduite proche de la coiffe en 5' de l'ARNm et n'est plus capable de rejoindre le codon de démarrage.

L'ajout des facteurs eIF1 et eIF1A à ce milieu permet au complexe ribosomal de se dissocier de la région 5' de l'ARNm, de débuter le balayage, et finalement conduit à la formation du complexe de démarrage au niveau du codon de démarrage de la traduction de l'ARNm.

Le complexe ribosomal 43S formé au niveau de la région 5' proche de la coiffe de l'ARNm est dans une position dite « ouverte » sur l'ARNm favorable au balayage de celui-ci jusqu'à la rencontre d'un codon de démarrage. L'ARNt initiateur dans ce complexe 43S compétant pour le balayage (Maag,

Algire, et Lorsch 2006), ne se trouve pas encore apparié au codon de démarrage. La conformation de l'ARNt a été appelée « P<sub>out</sub> » (Saini et al. 2009).

Dans ce complexe ouvert, le facteur eIF1 bloque la libération du Pi provenant de l'hydrolyse du GTP survenue dans le CT lors du balayage (Cheung et al. 2007; Algire, Maag, et Lorsch 2005).

#### 1.3.4.2 Les différentes régions d'eIF1A nécessaires à la reconnaissance du bon codon de démarrage de la traduction

La liaison d'eIF1 et d'eIF1A à la petite sous-unité du ribosome induit un changement de conformation de cette dernière, favorisant la liaison du complexe ternaire (CT) eIF2:GTP:Met-ARNt<sub>i</sub><sup>Met</sup> (Passmore et al. 2007; Mitchell et Lorsch 2008).

Cependant, seul le facteur eIF1A permet la stabilisation de l'interaction entre le CT et la petite sous-unité du ribosome (Lorsch et Dever 2010), retardant ainsi sa libération. En particulier, il a été montré que l'extrémité C-terminale d'eIF1A était impliquée dans la liaison du CT (Saini et al. 2009). Une paire de répétition dans l'extension C-terminale d'eIF1A, appelées régions activatrices du balayage (SE pour scanning enhancer) 1 et 2 (voir figure 1.23), sont identifiées comme stimulant le recrutement de l'ARNt initiateur lié à eIF2-GTP et empêchant le démarrage de la traduction sur un mauvais codon, par exemple UUG (Saini et al. 2009).

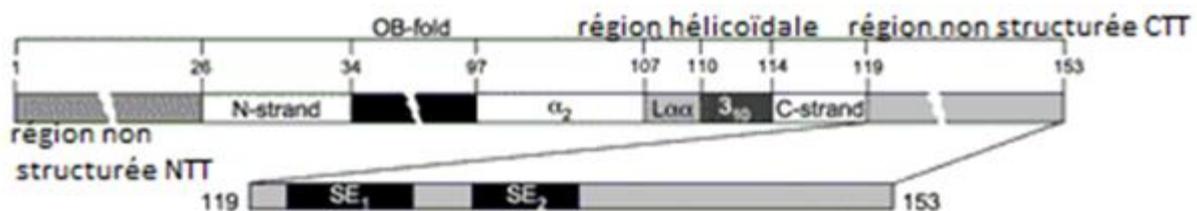


Figure 1.23 : carte des différents éléments composants eIF1A.

D'après (Saini et al. 2009)

De plus, des expériences de génétique mesurant le taux de traduction de l'ARNm GCN4 (Hinnebusch 1993) en utilisant une version tronquée dans son domaine C-terminal du facteur eIF1A ont montré l'importance de ce domaine dans la liaison des facteurs eIF2, eIF5 et eIF3 à la petite sous-unité du ribosome (Fekete et al. 2005). Cette région d'eIF1A pourrait aider à la stabilisation de l'ARNt<sub>i</sub><sup>Met</sup> sur la petite sous-unité du ribosome une fois le facteur eIF2-GDP libéré (Marintchev et al. 2003).

Les extrémités N-terminales et C-terminales d'eIF1A sont très proches du site P de liaison à l'ARNt initiateur, tout comme la boucle basique d'eIF1, ce qui empêcherait probablement ce dernier

d'entrer complètement dans le site P jusqu'à la reconnaissance du codon de démarrage, et aiderait ainsi à la sélection du bon codon de démarrage (Saini et al. 2009).

En opposition au domaine C-terminal, l'extension N-terminale du facteur eIF1A ainsi que le domaine hélicoïdal entravent le recrutement du CT et sont appelés éléments inhibiteur de balayage 1 et 2 (SI pour Scanning Inhibitor). Alors que les SE éléments vus précédemment améliorent la liaison du complexe ternaire dans une conformation permettant le balayage de l'ARNm mais pas le démarrage de la traduction, les SI éléments viendraient déstabiliser cette conformation, permettant ainsi l'accommodation complète de l'ARNt initiateur au site P lorsque le bon codon de démarrage est reconnu.

#### 1.3.4.3 Le départ d'eIF1 suite à l'appariement codon/anticodon

La modélisation d'un ARNt au niveau du site P du ribosome provenant de la structure cristallographique de la petite sous-unité 40S en complexe avec le facteur eIF1 (Rabl et al. 2011) met en évidence un choc stérique entre la boucle basique d'eIF1 et l'ARNt, suggérant que ce facteur pourrait être déplacé une fois l'appariement codon/anticodon établi. Dans ce contexte, la boucle basique d'eIF1 pourrait avoir un rôle de « senseur » de l'ARNt. En effet, le contact entre eIF1 et l'ARNt se situerait au niveau de la tige anticodon de l'ARNt, là où se trouvent les déterminants essentiels de l'ARNt initiateur, en particulier les trois paires GC. Ce contact pourrait expliquer comment eIF1 discrimine les ARNt initiateurs des ARNt élongateurs.

Lorsque l'appariement entre l'anticodon du Met-ARNt<sub>i</sub><sup>Met</sup> et le codon AUG de l'ARNm est réalisé, une réorganisation des facteurs du PIC se produit. Des expériences de FRET ont ainsi mis en évidence le déplacement du facteur eIF1 et de l'extrémité C-terminale d'eIF1A hors du site P du ribosome (Maag et al. 2005), permettant ainsi l'accommodation complète de l'ARNt initiateur au site P. Ce changement de conformation permettrait ensuite la dissociation d'eIF1 du complexe, conduisant à la libération du P<sub>i</sub> provenant de l'hydrolyse du GTP du complexe ternaire (Algire, Maag, et Lorsch 2005; Maag et al. 2005). De plus, le complexe pourrait alors passer d'un état ouvert permettant le balayage de l'ARNm dans lequel l'ARNt serait dans une conformation « P<sub>out</sub> », à un état fermé du ribosome où l'ARNt est apparié au codon de démarrage (conformation alors appelée « P<sub>in</sub> ») (Saini et al. 2009) permettant de début de l'étape d'allongement de la traduction.

Ainsi, au cours du démarrage de la traduction, eIF1 et eIF1A auraient des fonctions antagonistes conduisant finalement à la sélection du codon de démarrage AUG. eIF1 pourrait prévenir la formation prématurée de complexes à des codons qui ne seraient pas des codons de démarrage, et eIF1A pourrait quant à lui faciliter la stabilisation du complexe de démarrage sur le bon codon AUG.

## **1.4 Travail de thèse**

Avant le début de mon travail, les données structurales et fonctionnelles concernant les facteurs archées aIF1 et aIF1A étaient très peu nombreuses. Seulement deux études avaient été publiées par le groupe d'Udo Bläsi. En utilisant le système de l'archée *Sulfolobus solfataricus*, ce groupe montre qu'aIF1 se lie au ribosome pendant le démarrage de la traduction et que ce facteur favorise la liaison du complexe ternaire aIF2:GTP:Met-ARNT<sub>i</sub><sup>Met</sup> et de l'ARNm au ribosome (Hasenöhrl et al. 2006; Hasenöhrl et al. 2009).

Malgré de grandes similitudes entre les acteurs du démarrage de la traduction archée et eucaryote, l'étude du rôle des facteurs de démarrage archée garde tout son sens. En effet, malgré les avancées biochimiques permettant une meilleure description du complexe de démarrage de la traduction des eucaryotes, les mécanismes moléculaires expliquant la sélection du bon codon de démarrage par ce complexe restaient peu renseignés dans ce contexte. Le système de démarrage archée pouvant paraître comme un système simplifié du démarrage de la traduction eucaryote, l'étude du démarrage de la traduction chez les archées prend ici tout son sens.

De plus, certaines caractéristiques propres au démarrage archée, comme la présence d'une séquence SD en amont du codon de démarrage ou encore intrinsèque aux facteurs de démarrage, laissent penser que des mécanismes spécifiques aux archées existent et pourraient avoir été délaissés lors de l'évolution de la branche eucaryote.

Mon travail de thèse s'est inscrit dans la thématique principale de l'équipe concernant l'étude des mécanismes de démarrage de la traduction chez les archées.

Comme nous venons de le voir au cours de cette introduction, chez les archées, le balayage sur une longue distance de l'ARNm n'existe pas, compte tenu de la présence d'une séquence Shine-Dalgarno ou de très courtes régions 5' non traduites. Cependant, le démarrage de la traduction des archées utilise 3 facteurs de démarrage, aIF1, aIF1A et aIF2 homologues aux facteurs correspondants chez les eucaryotes. Ceci suggère que les facteurs archéens remplissent des fonctions identiques à celles de leurs homologues eucaryotes. Un tel parallèle est basé sur l'idée que la détection du codon de

démarrage chez les archées serait accomplie à la suite d'un balayage local du messenger par le complexe ribosomal, au sein d'une région du messenger compatible avec la maintenance de l'appariement SD/aSD.

Cette idée nous a conduits à la définition d'un cœur structural commun aux eucaryotes et aux archées, contrôlant la fidélité du processus de démarrage. Ce cœur structural est formé de la petite sous-unité du ribosome, de l'ARNm, de l'ARNt<sub>i</sub><sup>Met</sup> et des trois facteurs de démarrage a/eIF1, a/eIF1A et a/eIF2.

Le but de nos études a été d'appréhender le fonctionnement de ce cœur structural et d'identifier le rôle de chacun des facteurs le constituant.

#### **1.4.1 Construction d'une souche d'*E. coli* permettant la production d'ARNt initiateur d'intérêt**

La réalisation de ces études à tout d'abord nécessitait la mise en place d'un système permettant la reconstitution *in vitro* des complexes de démarrages archéens.

Le ribosome archéen ainsi que les trois facteurs de démarrage de la traduction archéens sont disponibles au laboratoire. Cependant, nous ne disposons pas d'ARNt initiateur archéen.

Lors de précédentes études menées au laboratoire, l'ARNt d'*E. coli* a été utilisé car le laboratoire est capable de le produire en grande quantité et de le purifier très efficacement. De plus, on sait aussi que cet ARN de transfert est lié correctement par le facteur aIF2 (Kd de l'ordre du nanomolaire). Malgré tout, cet ARNt<sup>fmet</sup> d'*E. coli* ne correspond pas au substrat naturel des archées. Afin de reproduire le plus fidèlement possible le complexe archéen, la production d'ARNt initiateur mimant un ARNt d'archée et pouvant être obtenu très pur et en grande quantité était nécessaire.

Dans ce but, j'ai tout d'abord construit une souche d'*E. coli* utilisant comme seule source d'ARNt initiateur un ARNt initiateur portant la mutation A<sub>1</sub>U<sub>72</sub>. En effet, comme nous avons pu le constater dans l'introduction, les différences de séquences entre un ARNt initiateur d'archée et un ARNt initiateur bactérien sont faibles. La différence la plus notable correspondant au remplacement des bases non appariées C<sub>1</sub>A<sub>72</sub> par les bases appariées A<sub>1</sub>U<sub>72</sub> dans l'ARNt initiateur archéen. Cette souche nous permet d'obtenir un ARNt initiateur mimant ceux des archées avec une très grande pureté. J'ai également pu déterminer la structure cristallographique de ce variant d'ARNt.

#### **1.4.2 Etude structurale du complexe de démarrage archée**

J'ai aussi pu contribuer à l'étude structurale par Cryo-microscopie électronique du complexe de démarrage de la traduction chez les archées. En effet, j'ai réalisé de nombreuses purifications de sous-unités ribosomales, de facteurs de démarrage ainsi que des variants archéens d'ARNt initiateur.

#### **1.4.3 Etude du rôle du facteur aIF1 lors du démarrage de la traduction**

Pour finir, l'essentiel de mon travail de thèse a porté sur l'étude du facteur de démarrage de la traduction aIF1 et dans une moindre mesure celui du facteur aIF1A. En effet, comme nous avons pu le constater lors de la présentation de ces facteurs, l'homologie des facteurs archéens avec les facteurs eucaryotes n'est pas totale, ce qui implique de possibles variations quant au rôle ou au mode de fonctionnement de ces derniers. Ainsi, l'étude de leur rôle lors du démarrage de la traduction chez les archées reste nécessaire pour la compréhension de ce mécanisme cellulaire essentiel dans les trois domaines du vivant.

Les récentes résolutions cristallographiques de la structure d'eIF1 ont permis de mettre en évidence des résidus pouvant jouer un rôle dans les fonctions de ce facteur.

Afin de pouvoir déterminer les résidus d'aIF1 impliqués dans la liaison de ce facteur au ribosome, j'ai mis au point une méthode de mesure utilisant l'anisotropie de fluorescence.

De plus, la technique du toeprinting a été employée afin de pouvoir observer le rôle d'aIF1 et de variants dans la formation des complexes de démarrage, et d'étudier l'influence des facteurs dans la formation de ces complexes.



## 2 Etude structurale de l'ARNt<sup>Met</sup><sub>f</sub> A<sub>1</sub>-U<sub>72</sub>

## 2.1 Introduction

Il existe chez la souche K-12 de la bactérie *Escherichia coli* deux types d'ARN<sub>f</sub><sup>Met</sup> (Mandal et Rajbhandary 1992) : l'ARN<sub>f</sub><sup>Met</sup><sub>1</sub>, espèce majoritaire codée par les 3 gènes redondants *metZ*, *metW* et *metV*, qui présente une 7-méthyl-guanine en position 46 et l'ARN<sub>f</sub><sup>Met</sup><sub>2</sub>, espèce minoritaire codée par le gène *metY*, qui présente une adénine à cette même position 46. Les ARNt initiateurs bactériens ne possèdent pas la paire de bases A<sub>1</sub>-U<sub>72</sub> de la tige acceptrice caractéristique des ARNt initiateurs eucaryotes et archées. Cette paire de base est importante dans la discrimination des ARNt initiateurs et élongateurs. De manière intéressante, le changement C<sub>1</sub>-A<sub>72</sub> de l'ARN<sub>f</sub><sup>Met</sup> d'*E. coli* en A<sub>1</sub>-U<sub>72</sub> permet d'obtenir un ARN de transfert dont la séquence de la tige acceptrice correspond au consensus de la séquence de la tige acceptrice des ARNt initiateurs d'archées (voir figure 2.1 ci-dessous).

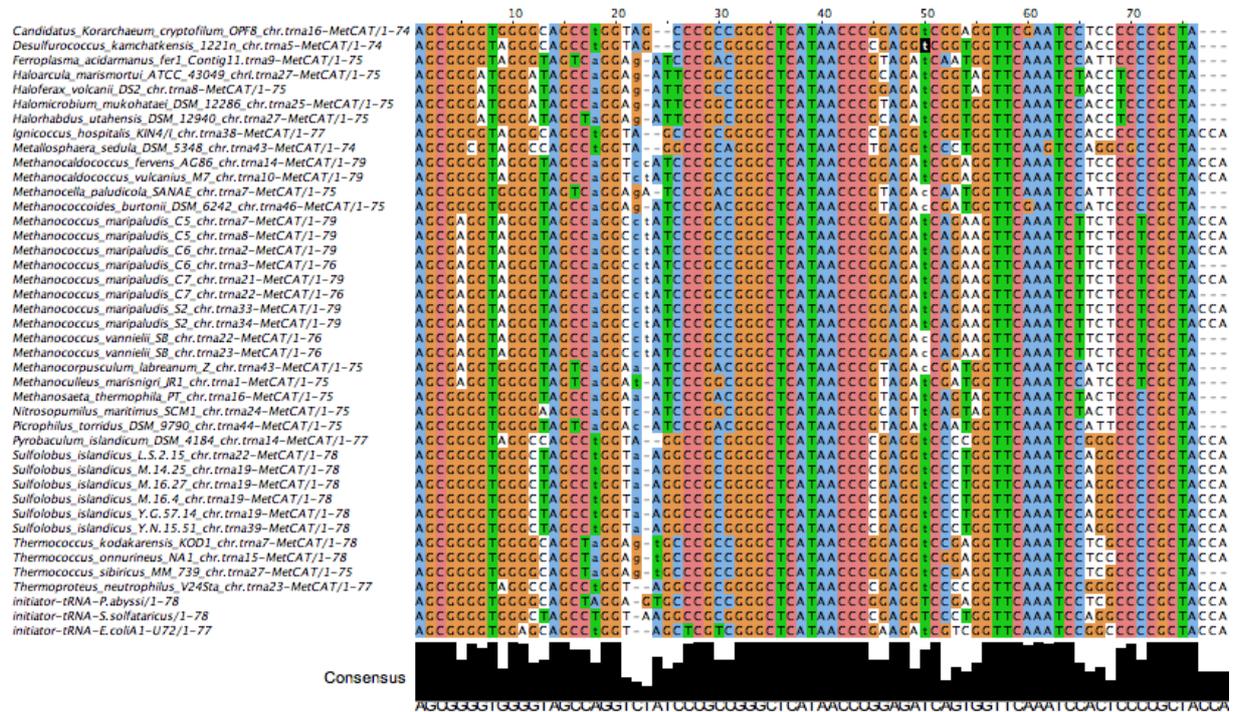


Figure 2.1 : alignement de séquence d'ARNt initiateurs d'archées

Ainsi, au laboratoire, nous avons été capables de produire en grande quantité des ARN<sub>f</sub><sup>Met</sup> portant la mutation A<sub>1</sub>-U<sub>72</sub> afin d'obtenir un bon mime d'un ARN<sub>t</sub> archée, et de pouvoir l'utiliser dans nos études structurales.

Afin d'inactiver les gènes *metZ*, *metW*, *metV* et *metY*, la technique d'inactivation de gènes chromosomiques par lambda red a été utilisée. Cette méthode, mise au point par Kirill A. Datsenko et Barry L. Wanner (Datsenko et Wanner 2000) permet d'inactiver des gènes par recombinaison

homologue en transformant la bactérie *Escherichia coli* avec un produit de PCR ayant une homologie avec le gène cible. Cette technique requiert la présence préalable de 3 gènes du phage Lambda portés par un plasmide :  $\gamma$ ,  $\beta$  et *exo*, codant respectivement pour les protéines Gam qui inhibe l'exonucléase V RecBCD d'*E. coli*, Bet et Exo qui permettent l'initiation de la recombinaison. En effet, l'exonucléase de type V Rec BCD présente chez *E. coli* doit être inactivée car elle dégrade les ADN linéaires (comme les produits de PCR) et limite ainsi les phénomènes de recombinaison.

Pour pouvoir sélectionner les clones inactivés dans le gène d'intérêt, une cassette portant un gène de résistance à un antibiotique (kanamycine) est utilisée pour la recombinaison. De plus, des séquences appelées FRT sont présentes de part et d'autre du gène de résistance afin de pouvoir par la suite éliminer cette cassette par une nouvelle recombinaison.

Dans un premier temps, la souche K-12 d'*E. coli* est inactivée des gènes codant pour l'ARNt initiateur 1. Une fois l'inactivation obtenue, la cassette de résistance à la kanamycine est éliminée grâce à l'activité de la protéine FLP recombinase qui reconnaît les sites FRT flanquant *kanR* et permet ainsi son élimination du génome d'*E. coli*. Les souches obtenues sont par la suite inactivées de leur second ARNt initiateur de la même façon que pour le/les premiers gènes. De plus, pour que cette seconde inactivation ne soit pas létale pour la cellule, celle-ci sera complétée par un plasmide contenant le gène codant pour un ARNt initiateur d'intérêt pour le laboratoire.

Une fois cette souche obtenue, nous avons produit et purifié en grande quantité cet ARNt<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> afin de réaliser des études structurales par cristallographie aux rayons X des reconstitutions de complexes de démarrage archée.

The structure of an *E. coli* tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> variant shows an unusual conformation of the A<sub>1</sub>-U<sub>72</sub> base pair

Auriane Monestier<sup>1</sup>, Pierre-Damien Coureux<sup>1</sup>, Michel Panvert<sup>1</sup>, Yves Mechulam<sup>1</sup> and  
Emmanuelle Schmitt<sup>1\*</sup>

<sup>1</sup>Laboratoire de Biochimie, Ecole polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau cedex,  
France

\*Corresponding author: emmanuelle.schmitt@polytechnique.edu

Laboratoire de Biochimie, Ecole Polytechnique, CNRS,

F-91128 Palaiseau cedex, FRANCE

Phone: + 33 1 69334885

Fax: + 33 1 69334909

**Running title:** *E. coli* tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> crystal structure

**En révision**

## ABSTRACT

Translation initiation in eukaryotes and archaea involves a methionylated initiator tRNA delivered to the ribosome in a ternary complex with e/aiF2 and GTP. Eukaryotic and archaeal initiator tRNAs contain a highly conserved A<sub>1</sub>-U<sub>72</sub> base pair at the top of the acceptor stem. The importance of this base pair to discriminate initiator tRNAs from elongator tRNAs has been established previously using genetics and biochemistry. However, no structural data illustrating how the A<sub>1</sub>-U<sub>72</sub> base pair participates in the accurate selection of the initiator tRNAs by the translation initiation systems are available. In this study, we constructed an *E. coli* strain using any initiator tRNA<sup>Met</sup> variant as the sole source of initiator tRNA, provided that the overproduced variant is able to sustain *E. coli* growth. Using such a strain, we purified large amounts of tRNA<sup>Met</sup><sub>f</sub> A<sub>1</sub>-U<sub>72</sub> devoid of wild-type tRNA<sup>Met</sup><sub>f</sub> contamination. The crystal structure of tRNA<sup>Met</sup><sub>f</sub> A<sub>1</sub>-U<sub>72</sub> was determined to 2.8 Å resolution. Interestingly, an unusual geometry of the A<sub>1</sub>-U<sub>72</sub> pair is observed. A possible role of this peculiar structure in the recognition of the initiator tRNA by its partners during eukaryotic and archaeal translation initiation is discussed.

## INTRODUCTION

The initiator tRNA<sub>i</sub><sup>Met</sup> (called tRNA<sub>f</sub><sup>Met</sup> in bacteria and organelles) plays a central role in decoding mRNA into proteins. Its recruitment by the translation initiation machinery allows setting of the reading frame of the mRNA by decoding the correct initiation codon. Selection of the initiator tRNA by the translation initiation machinery and its rejection by the elongation system involve some nucleotidic sequence features that distinguish it from elongator tRNAs, in particular elongator Met-tRNAs (Marck and Grosjean 2002; Kowitz and Lorsch 2010). Among them, whatever the domain of life, tRNA<sub>i</sub><sup>Met</sup> harbors a weak base pair or unpaired bases at position 1-72 whereas most elongator tRNAs

possess a strong G-C base pair at this position. In bacteria, a conserved mismatch at position 1:72 (most frequently C<sub>1</sub>:A<sub>72</sub>) is involved in the recognition of Met-tRNA<sub>f</sub><sup>Met</sup> by methionyl-tRNA<sub>f</sub><sup>Met</sup> transformylase, the enzyme responsible for the formylation of the esterified methionine. However, biochemical studies have shown that a weak base pair was tolerated by the enzyme (Guillon et al. 1992). Indeed, the specific interaction of methionyl-tRNA<sub>f</sub><sup>Met</sup> transformylase with a methionylated initiator tRNA is based on the ability of the enzyme to open the 1-72 pair (Schmitt et al. 1998). Moreover, the absence of strong base pairing at position 1-72 of the initiator tRNA was shown to be important for the rejection of the initiator tRNA by the elongation factor EF-Tu (Lee et al. 1991; Guillon et al. 1992; Rudinger et al. 1996) as well as for the protection of formyl-Met-tRNA<sub>f</sub><sup>Met</sup> against hydrolysis by peptidyl-tRNA hydrolase (Dutka et al. 1993).

In eukaryotes and in archaea, Met-tRNA<sub>i</sub><sup>Met</sup> binds to the small ribosomal subunit as a complex with the heterotrimeric factor e/aIF2 in its GTP-bound form. When pairing between AUG codon and CAU anticodon is correct, e/aIF2 is released from the ribosome in a GDP-bound state. Therefore, the selection of the start codon is achieved through the control of the nucleotide state of e/aIF2. Eukaryotic and archaeal initiator tRNAs almost systematically carry an A<sub>1</sub>-U<sub>72</sub> base pair (one exception in *Schizosaccharomyces pombe* tRNA<sub>i</sub><sup>Met</sup> carrying a Ψ<sub>1</sub>-A<sub>72</sub> base pair) (Keith et al. 1993; Marck and Grosjean 2002). This base pair has been shown crucial for the initiation process in various studies (von Pawel-Rammingen et al. 1992; Farruggio et al. 1996; Kapp and Lorsch 2004; Yatime et al. 2004; Kapp et al. 2006). By analogy with the interaction of methionyl-tRNA<sub>f</sub><sup>Met</sup> transformylase with Met-tRNA<sub>f</sub><sup>Met</sup> in bacteria, it was proposed that the opening of the 1-72 pair was important to ensure correct binding of the tRNA<sub>i</sub><sup>Met</sup> to e/aIF2 (Yatime et al. 2004). However, by studying Met-tRNA<sub>i</sub><sup>Met</sup> affinity for *S. cerevisiae* eIF2-GTP or eIF2-GDP, it was proposed that the identity and/or geometry of the 1-72 base pair was important for positioning the methionine moiety of the initiator tRNA in its recognition pocket and that an easily disrupted base pair may not be sufficient to insure full binding specificity (Kapp and Lorsch 2004). Finally, beyond its function in e/aIF2 recognition, the A<sub>1</sub>-U<sub>72</sub> base pair also influences the binding of the ternary complex to the 40S subunit and later steps in

translation initiation (Kapp et al. 2006). The available structure of the archaeal TC revealed the positioning of the initiator tRNA with respect to aIF2 but the resolution (5 Å) was too low to reveal the details of atomic interactions (Schmitt et al. 2012).

Structural studies of e/aIF2:tRNA, and more generally of eukaryotic and archaeal translation initiation complexes, would benefit from the availability of large amount of pure initiator tRNA carrying an A<sub>1</sub>-U<sub>72</sub> base pair. In this view, comparison of the sequence of archaeal *S. solfataricus* tRNA<sub>i</sub><sup>Met</sup> with that of *E. coli* tRNA<sub>f</sub><sup>Met</sup> shows 78% of identities between the two species (Figure 1A) with in particular identical acceptor stem, except that the archeal tRNA carries an A<sub>1</sub>-U<sub>72</sub> base pair whereas the bacterial tRNA has C<sub>1</sub>:A<sub>72</sub>. Therefore, the *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> variant appears as a good mimic of an archaeal initiator tRNA<sub>i</sub><sup>Met</sup>. To obtain pure fractions of such a tRNA, we designed an *E. coli* strain into which the four genes encoding initiator tRNA, *metZWW* and *metY* are inactivated and complemented by a plasmid borne gene coding for *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> variant as the sole source of initiator tRNA. We purified from this new strain large amounts of tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> devoid of wild-type tRNA<sub>f</sub><sup>Met</sup>. This mutant tRNA was crystallized and the structure was solved and refined to 2.8 Å resolution. Interestingly, the crystal structure shows a peculiar arrangement of the A<sub>1</sub>-U<sub>72</sub> pair. A possible role of the peculiar geometry of A<sub>1</sub>-U<sub>72</sub> base pair in the selection of the initiator tRNA by eukaryotic and archaeal translation initiation systems is discussed.

## RESULTS

### Construction of a strain for tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> expression

A strain using pBSTNAV-tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> as its sole source of initiator tRNA was constructed in three steps: knock-down of *metZWW*, transformation by pBSTNAV-tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> and knock-down of *metY*. First, the *metZWW* genes were deleted in *E. coli* MG1655 strain (Guyer et al. 1981). For this purpose, we have used the modified λRed/FLP procedure (Datsenko and Wanner 2000; Chen et al. 2015). The

whole knock-out procedure involved  $\lambda$ Red recombinase-mediated replacement of the *metZ WV* region by a *kan* marker flanked by two FRT sequences, one of which is mutated (Senecoff et al. 1988). The *kan* marker was then eliminated through FLP recombinase mediated reaction. The advantage of this procedure is that the mutant FRT sequence remaining on the chromosome at the knocked-out locus cannot react with any other FRT during subsequent modifications (Chen et al. 2015). Because of this and of the removal of the *kan* marker, the whole procedure can be repeated for knocking-out other genes. Recombination events were verified using PCR (Figure 1B). The resulting strain Pal $\Delta$ *metZ WV* could grow because initiator tRNA was produced from the intact *metY* locus, as previously observed (Kenri et al. 1991 ; Kapoor et al. 2011; Samhita et al. 2012). However, it displayed a slow growth phenotype, yielding small colonies, in agreement with previous studies of a  $\Delta$ *metZ WV::kan* strain (Samhita et al. 2014) (Figure 1C). In contrast, when we deleted *metY* using the same procedure, normal growth of the obtained Pal $\Delta$ *metY* strain was observed.

In order to knock-out the *metY* gene in Pal $\Delta$ *metZ WV*, cells were transformed by the  $\lambda$ Red-producing plasmid pKD46 (Datsenko and Wanner 2000) and then with pBSTNAV-tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub>. Introduction of the latter plasmid restored standard colony size, suggesting that the overproduced tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> was functional in translation initiation (Figure 1C). Interestingly, standard colony sizes were also restored with pBSTNAV-*S. solfataricus*-tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> (Figure 1C). Accordingly, the *metY* locus was successfully knocked-out through its  $\lambda$ Red-mediated replacement by the *kan* gene. The obtained *E. coli* strain, Pal $\Delta$ *metZ WV* $\Delta$ *metY::kan*:pBSTNAV-tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub>, has tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> as its sole source of initiator tRNA.

### **Overall structure of tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> and crystal packing analysis**

The strain described above was then used to overexpress and purify tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> as described in the Materials and Methods section. After dilution in buffer A, pure Met-tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> was used in

crystallization trials. Crystals belonging to space group C222<sub>1</sub> were obtained using 20% PEG 6000 and 2M NaCl as precipitants. The structure was solved by molecular replacement and refined to 2.8 Å resolution.

An analysis using DSSR (Lu et al. 2015) identifies all the secondary structure elements characteristic of the classical cloverleaf secondary structure as well as usual tertiary interactions that stabilize the L-shaped tertiary fold of the molecule. A global superimposition of the present structure to that of wild-type *E. coli* tRNA<sup>Met</sup> (PDB entry 3CW5; Barraud et al. 2008) shows modification in the angle made by the two branches of the tRNA, *i.e* the anticodon helix (from P<sub>56</sub> to P<sub>35</sub>) and the acceptor helix (from P<sub>56</sub> to P<sub>73</sub>) as defined (Giege 2008). The angle varies from 71° for *E. coli* tRNA<sup>Met</sup> to 81° for tRNA<sup>Met</sup><sub>A<sub>1</sub>-U<sub>72</sub></sub>. This difference reflects tRNA plasticity. Crystals containing tRNA<sup>Met</sup><sub>A<sub>1</sub>-U<sub>72</sub></sub> have been obtained in space group C222<sub>1</sub> whereas previous crystals of wild type *E. coli* tRNA<sup>Met</sup> had been obtained in space group P6<sub>4</sub>22. In the present crystals, three regions of the tRNA are involved in packing interactions (Figure 2). One contact point involves the anticodon loops of two tRNA molecules. Within this region, anticodon-anticodon base pairing between the C<sub>34</sub>A<sub>35</sub>U<sub>36</sub>A<sub>38</sub> nucleotides of one molecule and the corresponding nucleotides in a symmetry-related molecule is observed (Figure 2A). Two Watson-Crick A-U pairs and two non-canonical C-A hydrogen bonds are holding this interaction. Notably the same packing of anticodon loops was observed in P6<sub>4</sub>22 crystals of wild-type *E. coli* tRNA<sup>Met</sup> (Barraud et al. 2008).

A second contact point, specific to C222<sub>1</sub> crystals, involves the universal G<sub>19</sub>-C<sub>56</sub> base pair located at the elbow of the tRNA. The G<sub>19</sub>-C<sub>56</sub> base pair presents a flat hydrophobic surface stacked onto the corresponding base pair of a neighboring tRNA molecule (Figure 2B). Notably, this surface at the elbow is exploited for tRNA recognition by many partners including the ribosome (Korostelev et al. 2006; Zhang and Ferre-D'Amare 2013; Zhang and Ferre-D'Amare 2016).

Finally, a third contact point, specific to C222<sub>1</sub> crystals, exists between the top of the acceptor stem of one molecule and the minor groove of the T-stem of a neighbor. In particular, A<sub>73</sub> makes an A-

minor type I interaction (Nissen et al. 2001) with G<sub>51</sub>-C<sub>63</sub> receptor base pair as observed in a variant of group I self-splicing intron (Battle and Doudna 2002) (Figure 2C). A second interaction involves the NH<sub>2</sub> group of A<sub>1</sub> and the 2'OH group of a neighboring G<sub>64</sub>. Also, O1P of C<sub>74</sub> is bound to the NH<sub>2</sub> group of G<sub>52</sub>.

Because of crystal packing interactions, the tRNA molecules are more tightly bound together along the acceptor helices. This is reflected by lower B-values along the acceptor helix comparing with those observed for atoms belonging to the anticodon helix. According to this anisotropy, the TLS procedure was used in the last steps of refinement.

### **Conformation of the anticodon stem-loop**

Interestingly, the structure of the anticodon stem and loop is identical to that described for *E. coli* tRNA<sub>f</sub><sup>Met</sup> (Barraud et al. 2008). A<sub>37</sub> protrudes outside of the anticodon loop and A<sub>38</sub> is directly stacked onto U<sub>36</sub>. A C<sub>m32</sub>-A<sub>38</sub> unusual base pair is formed thereby extending the anticodon stem and restricting the anticodon loop to five nucleotides. The C<sub>m32</sub>-A<sub>38</sub> base pair is stabilized by a polar interaction between the N1 group of A<sub>38</sub> and the O2 group of C<sub>m32</sub> and by stacking onto G<sub>31</sub>-C<sub>39</sub>. Finally, A<sub>37</sub> is positioned within the major groove of the anticodon stem and possibly forms a base triple with G<sub>29</sub>-C<sub>41</sub> (Figure 3). Superimposition of tRNA<sub>f</sub><sup>Met</sup> and tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> structures shows that bendings of the anticodon loops slightly differ. However, despite this movement, the unusual network of interactions in the anticodon loop is conserved in the two tRNA structures.

### **An unusual A<sub>1</sub>-U<sub>72</sub> pair**

At the top of the acceptor stem of the tRNA, the A<sub>1</sub>-U<sub>72</sub> pair, characteristic of eukaryotic and archaeal initiator tRNAs, shows a peculiar conformation (Figure 4). Watson-Crick (WC) interactions between

A<sub>1</sub> and U<sub>72</sub> are unexpectedly not observed. As compared to a standard G<sub>1</sub>-C<sub>72</sub> WC base pair as in tRNA<sup>Phe</sup> (PDB entry 1EHZ (Shi and Moore 2000)), a rotation of base 1 around bond P-O3' is observed. Moreover, instead of adopting the usual *anti* conformation, the adenine ring rotates around the glycosidic bond and adopts a *syn* conformation ( $\chi=-14.9^\circ$ ). The 5' phosphoryl group folds back into the major groove of the acceptor stem and interacts with the N7 group of G2 (Figure 4A). On the other hand, base U<sub>72</sub> remains stacked onto C<sub>71</sub> as in a canonical tRNA<sup>Phe</sup> acceptor stem. As a consequence, the opening parameter (Lu et al. 2015) of the A<sub>1</sub>-U<sub>72</sub> base pair becomes unusually high (153.42°). The NH<sub>2</sub> group of A<sub>1</sub> points towards the minor groove of the acceptor helix. An interaction between the N1 group of A<sub>1</sub> and the O2 group of U<sub>72</sub> (d=3.0 Å) is predicted. In addition, A<sub>73</sub> is stacked onto A<sub>1</sub> instead of being stacked on the first base pair of the acceptor stem as in tRNA<sup>Phe</sup>. Finally, a water molecule bound to N7 of A<sub>73</sub> located just above U<sub>72</sub> could make an O- $\pi$  or an H- $\pi$  interaction with the uridine ring (Sarkhel et al. 2003). It is interesting to note that such water-nucleobase interactions would make up for the absence of base stacking on U<sub>72</sub> (Figure 4).

## DISCUSSION

The present work describes the construction of an *E. coli* strain able to grow in a genetic context where a variant of the initiator tRNA is used as the sole source of initiator tRNA. Using such a strain we observed that *E. coli* growth is supported by overexpression of an initiator tRNA carrying an A<sub>1</sub>-U<sub>72</sub> base pair instead of the natural C<sub>1</sub>:A<sub>72</sub>. This result is consistent with previous work showing that the absence of strong base pairing at position 1-72 is sufficient to insure efficient formylation of Met-tRNA<sub>f</sub><sup>Met</sup>, an important step for initiation of translation (Guillon et al. 1993; Mangroo and Rajbhandary 1995). Moreover, we also show that *E. coli* growth is supported by overexpression of the archaeal initiator tRNA<sub>i</sub><sup>Met</sup> from *S. solfataricus*. Accordingly, the sequence of the archaeal initiator tRNA<sub>i</sub><sup>Met</sup> from *S. solfataricus* is very close to that of an *E. coli* tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> variant. In particular, the sequences of the acceptor stems and of the T- and anticodon loops are identical in both tRNAs

(Figure 1A). Hence, the  $\lambda$ Red/FLP procedure applied on Pal $\Delta$ metZWV in the presence of an overproducing initiator tRNA plasmid gives a useful tool to test for initiator tRNA complementation and to allow overexpression and efficient purification of effective initiator tRNA variants in an *E. coli* context. This method represents a useful alternative to the P1 transduction strategy (Samhita et al. 2012) in particular in cases where it is desirable to limit the number of antibiotic resistance genes.

Using pure fractions of *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> variant, we obtained crystals and solved the 3D structure of the tRNA. The overall structure presents all the tertiary interactions responsible for the L-shaped folding of the tRNA as described (Barraud et al. 2008). Two regions exhibit specific structural features: the anticodon loop and the top of the acceptor stem. Within the anticodon loop, A<sub>37</sub> is extruded and A<sub>38</sub> makes a peculiar base pair with C<sub>32</sub>. Notably, the unusual conformation of the anticodon region was already observed in wild-type tRNA<sub>f</sub><sup>Met</sup> crystals obtained in a different space group (Barraud et al. 2008). In this view, it is notable that yeast initiator tRNA also shows an extruded base 37 (Basavappa and Sigler 1991). However, in this case, the presence of an N6-threonylcarbamoyladenosine modification in A<sub>37</sub> might explain this absence of a base triple interaction between t<sub>6</sub>A<sub>37</sub> and the G<sub>29</sub>-C<sub>41</sub> base pair. Together, these structures reflect the propensity of the anticodon loop of initiator tRNAs to adopt a '37-unstacked' conformation. This observation reinforces the idea that the a switch between a '37-unstacked' conformation observed in free-initiator tRNAs (present work and (Basavappa and Sigler 1991; Barraud et al. 2008)) as compared to the '37-stacked' conformation observed in tRNA<sub>f</sub><sup>Met</sup> bound to the P site (Selmer et al. 2006) might be important for initiator tRNA selection during the translation initiation process.

On another hand, tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> variant shows an unusual conformation of the A<sub>1</sub>-U<sub>72</sub> pair at the top of the acceptor stem. Notably, bases A<sub>1</sub> and U<sub>72</sub> are systematically encountered in archaeal initiator tRNAs and only one exception is known in eukaryotic initiator tRNAs (Marck and Grosjean 2002). Rather than the requirement of a weak base pair at position 1:72, recent biochemical results have suggested that the identity of the A<sub>1</sub>-U<sub>72</sub> base pair was important to insure the correct selection of

the initiator tRNA during translation initiation (reviewed in Kolitz and Lorsch 2010). In eukaryotes and in archaea, the initiator tRNA is delivered to the ribosome as a ternary complex, e/aIF2:GTP:tRNA (TC). In the TC, it was suggested that the geometry of the base pair at 1-72 was important for the correct positioning of the methionine moiety in its binding pocket (Kapp and Lorsch 2004). Furthermore, it was also proposed that the A<sub>1</sub>-U<sub>72</sub> base pair plays a role in the binding of the TC within the translation initiation complex and directly affects translation efficiency (Kapp et al. 2006).

Notably, the conformation of the anticodon loop as well as that of the A<sub>1</sub>-U<sub>72</sub> base pair are stabilized by unusual protonation schemes. Indeed, the C<sub>m32</sub>-A<sub>38</sub> pair involves a polar contact between the N<sub>1</sub> group of A<sub>38</sub> and the O<sub>2</sub> group of C<sub>m32</sub> (d=3.3 Å) suggesting that the N<sub>1</sub> group of A<sub>38</sub> is protonated. Likewise, interaction of N<sub>1</sub> group of A<sub>1</sub> with O<sub>2</sub> group of U<sub>72</sub> (d=3.0 Å) assumes that N<sub>1</sub> group of A<sub>1</sub> is protonated. Because crystals have been obtained at pH 6.3 and because the pKa of adenosine 5'phosphate is about 3.9 (Saenger 1984), it is likely that crystal packing interactions have influenced local changes of the pKa values for A<sub>1</sub> and A<sub>38</sub>. Adenine protonations with unusually elevated pKa have already been described for instance in domain B of hairpin ribozyme (Ravindranathan et al. 2000) or in tRNA<sup>Lys</sup> anticodon stem-loop (Durant and Davis 1999). Alternatively, it cannot be totally excluded that a protonation scheme compatible with a hydrogen bond between N<sub>1</sub> of A<sub>1</sub> and O<sub>2</sub> of U<sub>72</sub> occurs through tautomerization of the adenine in its imino form although large energetic penalties render this hypothesis less likely.

Whatever the mechanism, the crystalline form of tRNA<sup>Met</sup><sub>f</sub>-A<sub>1</sub>-U<sub>72</sub> is able to provide a favorable environment to stabilize the peculiar conformation of the A<sub>1</sub>-U<sub>72</sub> base pair. Even if stabilized by crystal packing, the conformation must exist at least transiently in solution to allow crystal nucleation and growth. Altogether, this renders possible that an environment favoring the unusual A<sub>1</sub>-U<sub>72</sub> base pair might also exist during interaction of the initiator tRNA with one of its partner during translation initiation. One obvious candidate is e/aIF2-GTP within the TC. In this view, the peculiar A<sub>1</sub>-U<sub>72</sub> conformation might explain its role in the positioning of the methionine moiety of the initiator tRNA

in its recognition pocket within the  $\gamma$  subunit of eIF2-GTP (Kapp and Lorsch 2004). Notably, the crystal structure of *S. cerevisiae* initiator tRNA, obtained in P6<sub>4</sub>22 space group, shows a standard WC A<sub>1</sub>-U<sub>72</sub> base pair (Basavappa and Sigler 1991). Therefore, the tRNA may switch between the two conformations depending on the stage its functional cycle. Unfortunately, the available 3D structure of archaeal aIF2:GDPNP:Met-tRNA<sub>f</sub><sup>Met</sup> was obtained at low resolution (5Å) thus preventing to observe the fine geometrical requirements for optimal tRNA binding (Schmitt et al. 2012). Therefore, further structural studies of the TC are required to reveal the conformation of the A<sub>1</sub>-U<sub>72</sub> base pair of the initiator tRNA bound to e/aIF2. Finally, recent Cryo-EM studies of archaeal translation initiation complexes (our laboratory, to be published) show deformation of the TC structure at some stages of the translation initiation process. In this view, one can also propose that the adjustments of the A<sub>1</sub>-U<sub>72</sub> base pair geometry participate in the control of the dynamics of the initiation complex during start codon recognition. To conclude, the present structure gives a new example of an RNA structure into which rare ionic forms or tautomeric states of bases could be at the origin of unusual base pairing important for the fidelity of the translational process (Rozov et al. 2015; Rozov et al. 2016).

## **MATERIALS and METHODS**

### **Inactivation of genes in the *E. coli* chromosome**

Gene deletions were performed using the  $\lambda$ Red/FLP procedure (Datsenko and Wanner 2000) as modified (Chen et al. 2015). *E. coli* K12 MG1655 strain (F<sup>-</sup>  $\lambda$  *ilvG* *rfb*-50 *rph*-1) was transformed with the pKD46 plasmid, carrying a thermosensitive replicon and the  $\lambda$  Red operon under the control of an arabinose inducible promoter. Cells were plated on LB plates supplemented with ampicillin (50  $\mu$ g/mL) and grown at 30°C. The *kan* gene flanked by FRT (Flp recombinase Recognition Target) sequences carried by pEAW507 was PCR-amplified using primers made of ca 50 bases immediately

adjacent to the region to be deleted fused to 19 bases bordering the FRT-kan-FRT cassette on the plasmid. Sequences of the oligonucleotides used for the deletions of *metZ<sub>WV</sub>* and *metY* were as follows :

**5'AAAAAAGGTTGCATGAAAACGCGAGCGGAGTATAGTGCGCATCCACGGACGTAGAGGATCCC  
GGAAGT3'** (*metZ<sub>WV</sub>*-forward)

**5'TTTATGGGAGACCAGAAACAAAAAACACCCGTTAGGGTGTTCAATAATTCCCTTTCGTCTTCAAG  
AAT3'** (*metZ<sub>WV</sub>*-reverse)

**5'GCAGTATTTGCATTTTTTACCCAAAACGAGTAGAATTTGCCACGTTTCAGGCGTAGAGGATCCCGG  
AAGT** (*metY*-forward)

**5'CGAAGGCCGAAGTCTTCACAGTATATTTGAAAAAGGACTCTAAGGGAAAGCCCTTTCGTCTTCAA  
GAAT3'** (*metY*-reverse)

The PCR product was digested by DpnI in order to degrade the plasmidic matrix, and transformed by electroporation into MG1655-pKD46 cells grown at 30°C in LB medium containing 1 mM L-arabinose. After electroporation, 1 mL of LB medium containing 1 mM L-arabinose was added and cells were incubated 2 hours at 30°C plus 4 hours at 37°C, and finally plated at 37°C on LB plates containing kanamycin (25 µg/mL). Clones were then tested by PCR for correct recombination of the cassette at the desired locus and for loss of the pKD46 plasmid. PCR primers were as follows:

5'GGCGCAACGAAGATAACAAACCGCCGC3' (ver-*metZ<sub>WV</sub>*-forward),

5'GTATTTTGCATGGGGCGACGTGGCGA3' (ver-*metZ<sub>WV</sub>*-reverse),

5'AGTATTCAACAAATGAAAGTGAAGTGG3' (ver-*metY*-forward)

and

5'TAACAAAAACCCCGATAAATCGGGGC3' (ver-*metY*-reverse).

For elimination of the *kan* gene, cells were first transformed with pLH29 (Huang et al. 1997), a plasmid expressing the FLP recombinase. One colony was grown overnight at 37 °C in 20 mL of LB containing 25 µg/mL chloramphenicol and 1 mM IPTG, and then left two days at room temperature. Clones were then screened for loss of resistance to kanamycin, and removal of the *kan* gene was then verified by PCR using the primers described above (Figure 1B). Finally, plasmid pLH29 was eliminated by performing four successive liquid cultures to saturation in LB at 37°C. Individual clones were finally tested for loss of pLH29 by assaying sensitivity to chloramphenicol. This procedure yielded PalΔmetZWV and PalΔmetY strains.

For inactivation of *metY* in PalΔmetZWV carrying pBSTNAV-tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> derivative, the same procedure was used except that pKD46 was selected with gentamycin instead of ampicillin. In this case, the *kan* gene at the *metY* locus was not removed after *metY* deletion. This procedure yielded PalΔmetZWVΔmetY::kan strains carrying a pBSTNAV- tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub>.

### **tRNA purification**

The *E. coli* tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> variant was overexpressed in PalΔmetZWVΔmetY::kan carrying pBSTNAV-tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> and purified as described (Mechulam et al. 2007). Briefly, a whole tRNA extract was prepared according to the Zubay procedure (Zubay 1962). The overexpressed tRNA<sup>Met</sup><sub>f</sub> was then separated from other tRNAs by an anion exchange step on a Q-Hiload column (16 x 100 mm) equilibrated in 20 mM Tris-HCl pH 7.6, 200 mM NaCl, 8 mM MgCl<sub>2</sub>, 0.1 mM EDTA. tRNAs were eluted using a 350 mM to 550 mM NaCl gradient in the same buffer. About 10 mg of tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> was obtained from 1L of *E. coli* culture. Methionine acceptance of the tRNA<sup>Met</sup><sub>f</sub>A<sub>1</sub>-U<sub>72</sub> preparation ranged between 1450 and 1600 pmol/A<sub>260</sub> unit. Before crystallization, the tRNA was methionylated and purified as described in (Mechulam et al. 2007). The aminoacylated tRNA was resuspended in water

(final concentration 400  $\mu\text{M}$ ) and diluted to 50  $\mu\text{M}$  in buffer A (10 mM MOPS pH 6.7, 5 mM  $\text{MgCl}_2$ , 200 mM NaCl, 10 mM 2-mercaptoethanol) before crystallization trials.

### **Crystallization, data collection, structure solution and refinement**

Crystals belonging to space group  $C222_1$  were obtained using the sitting drop method with a 1:1 mixture of a 50  $\mu\text{M}$  solution of Met-tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> in buffer A and 20% PEG 6000, 2M NaCl as precipitants. Final pH of the 1:1 mixture of buffer A and the precipitant was 6.3. Interestingly, the  $C222_1$  space group was never observed with crystals of wild type *E. coli* initiator tRNA<sub>f</sub><sup>Met</sup>. A highly redundant dataset was collected. According to the CC(1/2) criterion (Table 1) the high resolution limit was set to 2.8 Å (Karplus and Diederichs 2012; Diederichs and Karplus 2013). The structure was solved by molecular replacement using PHASER and the coordinates of *E. coli* tRNA<sub>f</sub><sup>Met</sup> (PDB entry 3CW5 (Barraud et al. 2008)) into which the terminal CCA bases have been removed. Initial map showed different orientations of the anticodon region and of the 3' extremity of the acceptor helix as compared to the molecular replacement-starting model. Cycles of manual reconstructions, performed with the program COOT (Emsley et al. 2010), and of refinements steps with PHENIX (Adams et al. 2010) led to a model refined to 2.8 Å resolution ( $R=0.196$ ,  $R_{\text{free}}=0.225$ ). Because crystal packing interactions are weaker in one direction, B factors were refined using TLS (Adams et al. 2010). All nucleotides are visible except A<sub>76</sub>. Moreover, positioning of C<sub>75</sub> is only tentative.

### **ACCESSION NUMBERS**

The Protein Data Bank accession number for the structure reported in this article is 5L4O.

### **ACKNOWLEDGEMENTS**

We thank the staff of the macromolecular crystallography beamlines PX1 and PX2 at the SOLEIL synchrotron (Saclay, France) for expert assistance during data collection. This work was supported by grants from the Centre National de la Recherche Scientifique and by Ecole polytechnique to Unité Mixte de Recherche n°7654. AM was a recipient of a Gaspard Monge PhD fellowship from Ecole polytechnique.

## REFERENCES

- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst* **D66**: 213-221.
- Barraud P, Schmitt E, Mechulam Y, Dardel F, Tisne C. 2008. A unique conformation of the anticodon stem-loop is associated with the capacity of tRNA<sup>Met</sup> to initiate protein synthesis. *Nucleic Acids Res* **36**: 4894-4901.
- Basavappa R, Sigler PB. 1991. The 3 Å crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination. *EMBO J* **10**: 3105-3111.
- Battle DJ, Doudna JA. 2002. Specificity of RNA-RNA helix recognition. *Proc Natl Acad Sci U S A* **99**: 11676-11681.
- Chen SH, Byrne RT, Wood EA, Cox MM. 2015. Escherichia coli radD (yejH) gene: a novel function involved in radiation resistance and double-strand break repair. *Mol Microbiol* **95**: 754-768.
- Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.
- Diederichs K, Karplus PA. 2013. Better models by discarding data? *Acta Crystallogr D Biol Crystallogr* **69**: 1215-1222.
- Durant PC, Davis DR. 1999. Stabilization of the anticodon stem-loop of tRNA<sup>Lys,3</sup> by an A+-C base-pair and by pseudouridine. *J Mol Biol* **285**: 115-131.
- Dutka S, Meinnel T, Lazennec C, Mechulam Y, Blanquet S. 1993. Role of the 1-72 base pair in tRNAs for the activity of *Escherichia coli* peptidyl-tRNA hydrolase. *Nucleic Acids Research* **21**: 4025-4030.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr* **D66**: 486-501.
- Farruggio D, Chaudhuri J, Maitra U, RajBhandary UL. 1996. The A1 x U72 base pair conserved in eukaryotic initiator tRNAs is important specifically for binding to the eukaryotic translation initiation factor eIF2. *Mol Cell Biol* **16**: 4248-4256.
- Giege R. 2008. Toward a more complete view of tRNA biology. *Nat Struct Mol Biol* **15**: 1007-1014.
- Guillon JM, Mechulam Y, Blanquet S, Fayat G. 1993. Importance of formylability and anticodon stem sequence to give tRNA<sup>Met</sup> an initiator identity in *Escherichia coli*. *J Bacteriol* **175**: 4507-4514.
- Guillon JM, Meinnel T, Mechulam Y, Lazennec C, Blanquet S, Fayat G. 1992. Nucleotides of tRNA governing the specificity of *Escherichia coli* methionyl-tRNA(fMet) formyltransferase. *J Mol Biol* **224**: 359-367.

- Guyer MS, Reed RR, Steitz JA, Low KB. 1981. Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spring Harb Symp Quant Biol* **45 Pt 1**: 135-140.
- Huang HK, Yoon H, Hannig EM, Donahue TF. 1997. GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. *Genes Dev* **11**: 2396-2413.
- Kapoor S, Das G, Varshney U. 2011. Crucial contribution of the multiple copies of the initiator tRNA genes in the fidelity of tRNA(fMet) selection on the ribosomal P-site in *Escherichia coli*. *Nucleic Acids Res* **39**: 202-212.
- Kapp LD, Kolitz SE, Lorsch JR. 2006. Yeast initiator tRNA identity elements cooperate to influence multiple steps of translation initiation. *RNA* **12**: 751-764.
- Kapp LD, Lorsch JR. 2004. GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2. *J Mol Biol* **335**: 923-936.
- Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science* **336**: 1030-1033.
- Keith G, Heitzler J, el Adlouni C, Glasser AL, Fix C, Desgres J, Dirheimer G. 1993. The primary structure of cytoplasmic initiator tRNA(Met) from *Schizosaccharomyces pombe*. *Nucleic Acids Res* **21**: 2949.
- Kenri T, Kohno K, Goshima N, Imamoto F, Kano Y. 1991. Construction and characterization of an *Escherichia coli* mutant with a deletion of the metZ gene encoding tRNA (f1Met). *Gene* **103**: 31-36.
- Kolitz SE, Lorsch JR. 2010. Eukaryotic initiator tRNA: finely tuned and ready for action. *FEBS Lett* **584**: 396-404.
- Korostelev A, Trakhanov S, Laurberg M, Noller HF. 2006. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell* **126**: 1065-1077.
- Lee CP, Seong BL, RalBhandary UL. 1991. Structural and sequence elements important for recognition of *Escherichia coli* formylmethionine tRNA by methionyl-tRNA transformylase are clustered in the acceptor stem. *J Biol Chem* **266**: 18012-18017.
- Lu XJ, Bussemaker HJ, Olson WK. 2015. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res* **43**: e142.
- Mangroo D, Rajbhandary UL. 1995. Mutants of *Escherichia coli* initiator tRNA defective in initiation. *J Biol Chem* **270**: 12203-12209.
- Marck C, Grosjean H. 2002. tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* **8**: 1189-1232.
- Mechulam Y, Guillon L, Yatime L, Blanquet S, Schmitt E. 2007. Protection-based assays to measure aminoacyl-tRNA binding to translation initiation factors. *Methods Enzymol* **430**: 265-281.

- Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. 2001. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc Natl Acad Sci U S A* **98**: 4899-4903.
- Ravindranathan S, Butcher SE, Feigon J. 2000. Adenine protonation in domain B of the hairpin ribozyme. *Biochemistry* **39**: 16026-16032.
- Rozov A, Demeshkina N, Khusainov I, Westhof E, Yusupov M, Yusupova G. 2016. Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat Commun* **7**: 10457.
- Rozov A, Demeshkina N, Westhof E, Yusupov M, Yusupova G. 2015. Structural insights into the translational infidelity mechanism. *Nat Commun* **6**: 7251.
- Rudinger J, Hillenbrandt R, Sprinzl M, Giege R. 1996. Antideterminants present in minihelix(Sec) hinder its recognition by prokaryotic elongation factor Tu. *EMBO J* **15**: 650-657.
- Saenger W. 1984. *Principles of Nucleic Acid Structure*. Springer-Verlag, New-York, Heidelberg, Berlin.
- Samhita L, Nanjundiah V, Varshney U. 2014. How many initiator tRNA genes does Escherichia coli need? *J Bacteriol* **196**: 2607-2615.
- Samhita L, Shetty S, Varshney U. 2012. Unconventional initiator tRNAs sustain Escherichia coli. *Proc Natl Acad Sci U S A* **109**: 13058-13063.
- Sarkhel S, Rich A, Egli M. 2003. Water-nucleobase "stacking": H-pi and lone pair-pi interactions in the atomic resolution crystal structure of an RNA pseudoknot. *J Am Chem Soc* **125**: 8998-8999.
- Schmitt E, Panvert M, Blanquet S, 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, Panvert M, Lazennec-Schurdevin C, Coureux PD, Perez J, Thompson A, Mechulam Y. 2012. Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat Struct Mol Biol* **19**: 450-454.
- Schrodinger, LLC. 2010. The PyMOL Molecular Graphics System, Version 1.3r1.
- Selmer M, Dunham CM, Murphy FVt, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**: 1935-1942.
- Senecoff JF, Rossmeissl PJ, Cox MM. 1988. DNA recognition by the FLP recombinase of the yeast 2 mu plasmid. A mutational analysis of the FLP binding site. *J Mol Biol* **201**: 405-421.
- Shi H, Moore PB. 2000. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA* **6**: 1091-1105.
- von Pawel-Rammingen U, Astrom S, Bystrom AS. 1992. Mutational analysis of conserved positions potentially important for initiator tRNA function in Saccharomyces cerevisiae. *Mol Cell Biol* **12**: 1432-1442.

- Yatime L, Schmitt E, Blanquet S, Mechulam Y. 2004. Functional molecular mapping of archaeal translation initiation factor 2. *J Biol Chem* **279**: 15984-15993.
- Zhang J, Ferre-D'Amare AR. 2013. Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* **500**: 363-366.
- . 2016. The tRNA Elbow in Structure, Recognition and Evolution. *Life* **6**.
- Zubay G. 1962. The isolation and fractionation of soluble ribonucleic acids. *J Mol Biol* **4**: 347-356.

TABLE

<b>Data Collection</b>	tRNA <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub>
Crystallization	10% PEG 6000
Conditions	2M NaCl
Space group	C222 <sub>1</sub>
Cell dimensions	
<i>a, b, c</i> (Å)	35.15 142.62 202.12
<i>α, β, γ</i> (°)	90 90 90
Resolution (Å)	41.2-2.8
<i>R</i> <sub>sym</sub> (%)	8.4 (246.4) <sup>a</sup>
<i>I</i> / <i>σ</i> <i>I</i>	17.03 (0.95)
CC(1/2) (%)	99.9 (74.4)
Completeness (%)	99.2 (95.5)
Redundancy	9.84 (9.74)
<b>Refinement</b>	
Resolution (Å)	35.2-2.8
No. reflections	13004
<i>R</i> <sub>work</sub> ( <i>R</i> <sub>free</sub> )	0.196 (0.225)
No. atoms/B-factors (Å <sup>2</sup> )	
tRNA	1623/116.7
Water	5/92.8
Na+	1/70.5
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	1.14

**Table 1:** Data collection and refinement statistics for tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> structure determination

A single crystal was used for data collection.

<sup>a</sup>Values in parentheses are for highest-resolution shell.

## Figure Legends

### Figure 1: Inactivation of *metZ WV* leads to a slow growth phenotype.

(A) Cloverleaf representations of initiator Met-tRNA<sup>Met</sup>. (Left) *S. solfataricus* Met-tRNA<sub>i</sub><sup>Met</sup>, (Right) *E. coli* Met-tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> variant. Note that wild-type tRNA<sub>f</sub><sup>Met</sup> carries C<sub>1</sub>:A<sub>72</sub>. Nucleotides in red in *S. solfataricus* Met-tRNA<sub>i</sub><sup>Met</sup> are not found in *E. coli* Met-tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub>. Post-transcriptional modifications of *S. solfataricus* Met-tRNA<sub>i</sub><sup>Met</sup> are not known.

(B) Control by PCR of knocked-down cells. The *metY* and *metZ* loci of each strain were probed by PCR analysis using whole cells as template DNA sources. Primer couples for probing the *metY* locus (VerY,

lanes 1,2,3) or the *metZ* locus (VerZ, lanes 4,5,6). Shown are the analyses of

PalΔ*metZ WV*Δ*metY*:kanR- pBSTNAV-tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> (lanes 1 and 4), PalΔ*metZ WV* (lanes 2 and 5) and

PalΔ*metY* (lanes 3 and 6). Expected sizes are 1411 bp

(lane 1), 406 bp (lane 2), 398 bp (lane 3), 364 bp (lanes 4 and 5) and 591 bp (lane 6). Lane M was loaded with a molecular weight marker (MWD-1, Nippongenetics) relevant band sizes of which are indicated on the left.

(C) Strain PalΔ*metZ WV* transformed with PBSTnav2 displays a slow growth phenotype (small colonies). The phenotype can be reversed upon expression of either tRNA<sub>f</sub><sup>Met</sup>, tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> or *S. solfataricus* tRNA<sub>i</sub><sup>Met</sup>.

### Figure 2: Crystal packing interactions.

The central view shows arrangement of the tRNA molecules within the C222<sub>1</sub> crystals with the three contact points boxed and detailed in the close-up views. The reference molecule of the asymmetric unit is in grey and the neighboring molecules are in orange. In the close-up views, one molecule is shown with bases in cyan and the neighboring molecule is in orange. Intramolecular interactions are

drawn with red dashed lines and intermolecular interactions are drawn with black dashed lines. Note that the close-up views are not drawn in the same orientations as the central view.

(A) Crystal packing at the anticodon loop.

(B) Crystal packing at the elbow.

(C) Crystal packing at the top of the acceptor stem. The left view shows  $A_{73}$  of a tRNA molecule making an A-minor interaction with the  $C_{51}$ - $G_{63}$  pair of a neighboring molecule. The right view shows interaction of the acceptor stem of a tRNA molecule with the T-stem of a crystallographic neighbor.

**Figure 3: Conformation of the anticodon loop of the *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> variant.**

The figure shows  $A_{37}$  outside of the anticodon loop forming an A-minor interaction with the  $G_{29}$ - $C_{41}$  base pair.  $A_{38}$  is stacked between  $U_{36}$  and  $C_{39}$  and forms an unusual base pair with  $C_{32}$ .

**Figure 4: Conformation of the top of the acceptor stem of the *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> variant.**

(A) Views in two orientations of the 2.8 Å resolution '2mFo-DFc' map of *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> contoured at 1.4 standard deviation in the region of the A<sub>1</sub>-U<sub>72</sub> base pair.

(B) The views in two different orientations compare the geometry of a standard  $G_1$ - $C_{72}$  base pair as observed in tRNA<sup>Phe</sup> (yellow sticks, PDB entry 1EHZ; Shi and Moore 2000) with that of the A<sub>1</sub>-U<sub>72</sub> base pair of *E. coli* tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> (cyan sticks). Comparison results from the superimposition of the acceptor stems of the two tRNAs.

The figures were drawn with Pymol (Schrodinger 2010)

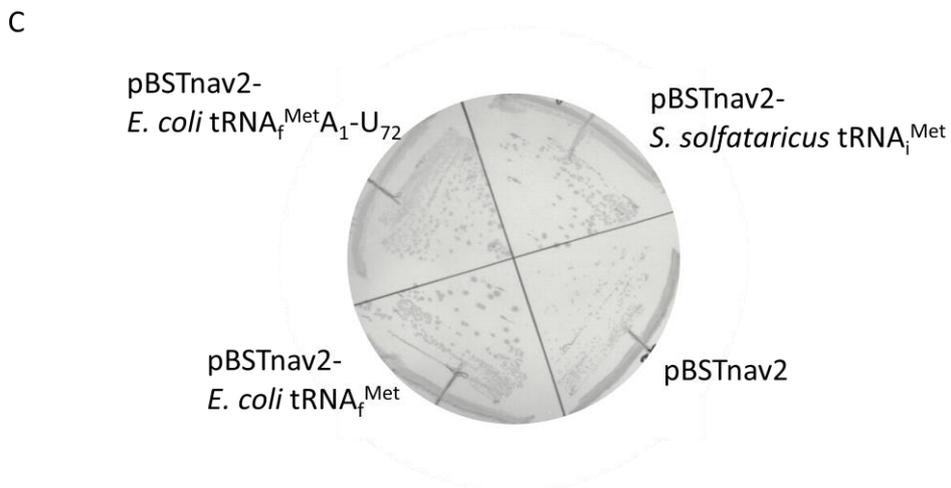
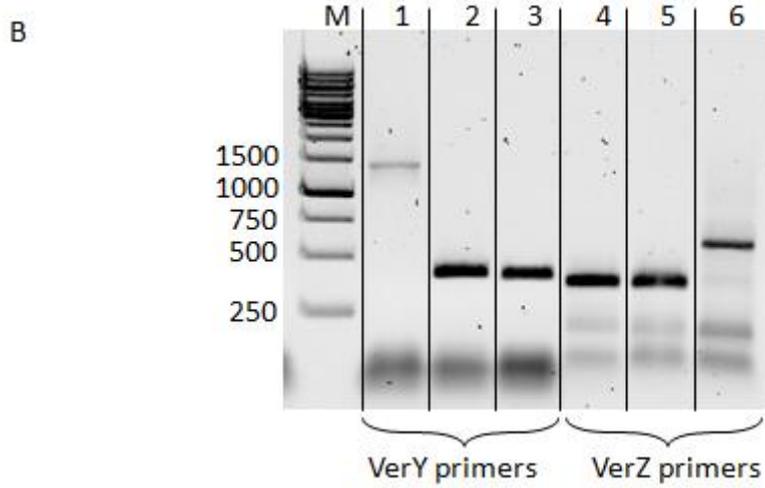
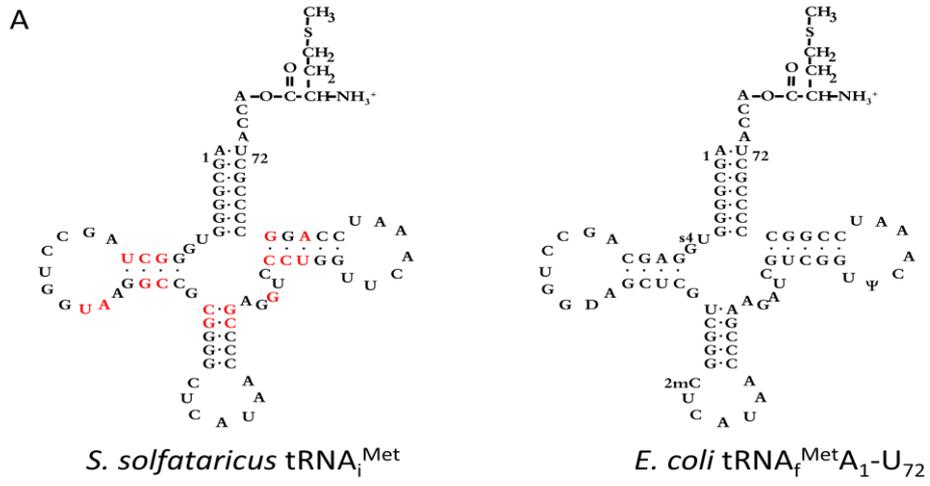


Figure 1

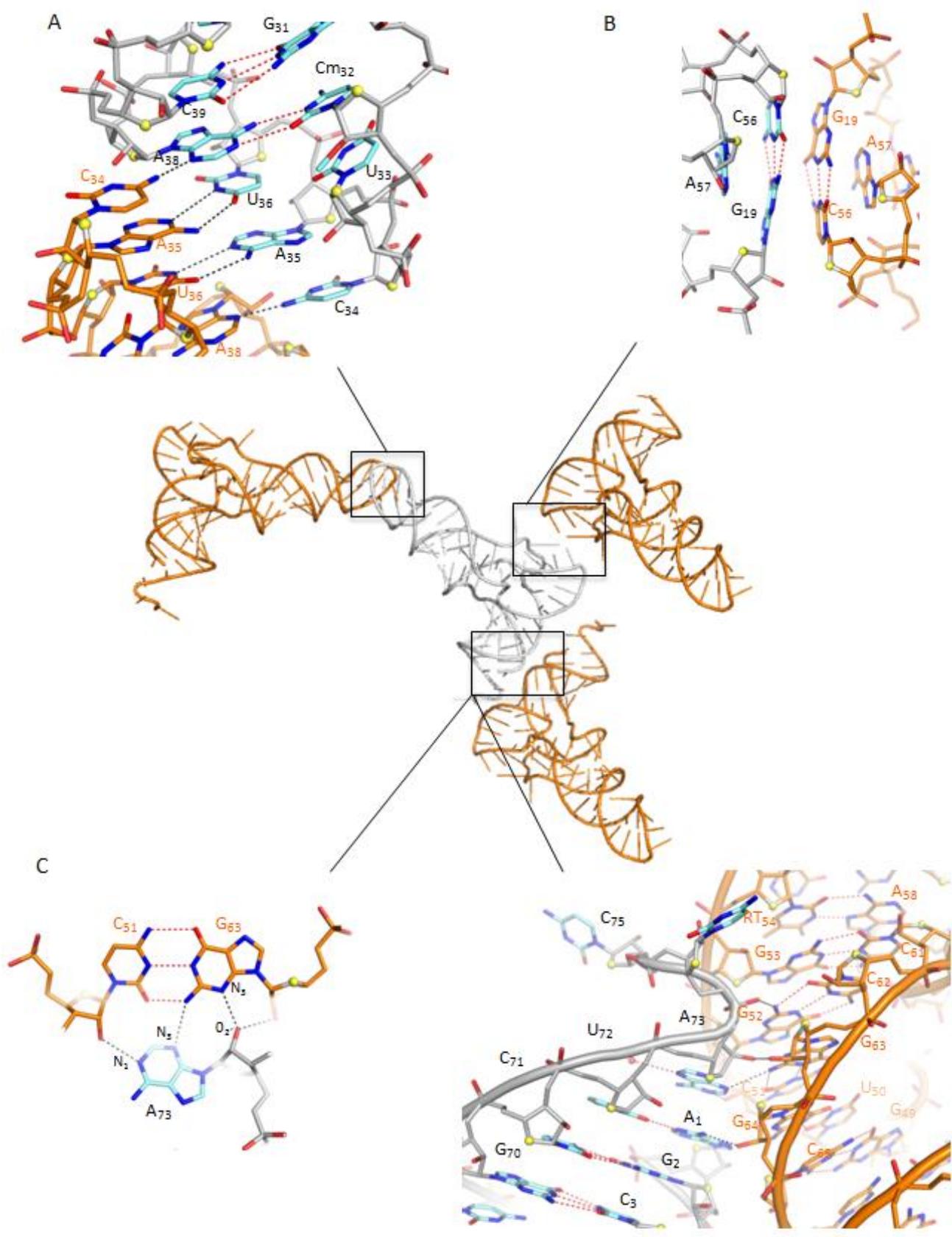


Figure 2

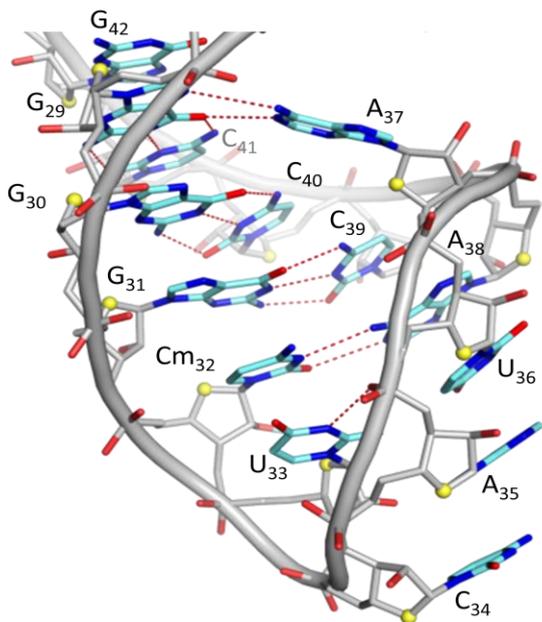


Figure 3

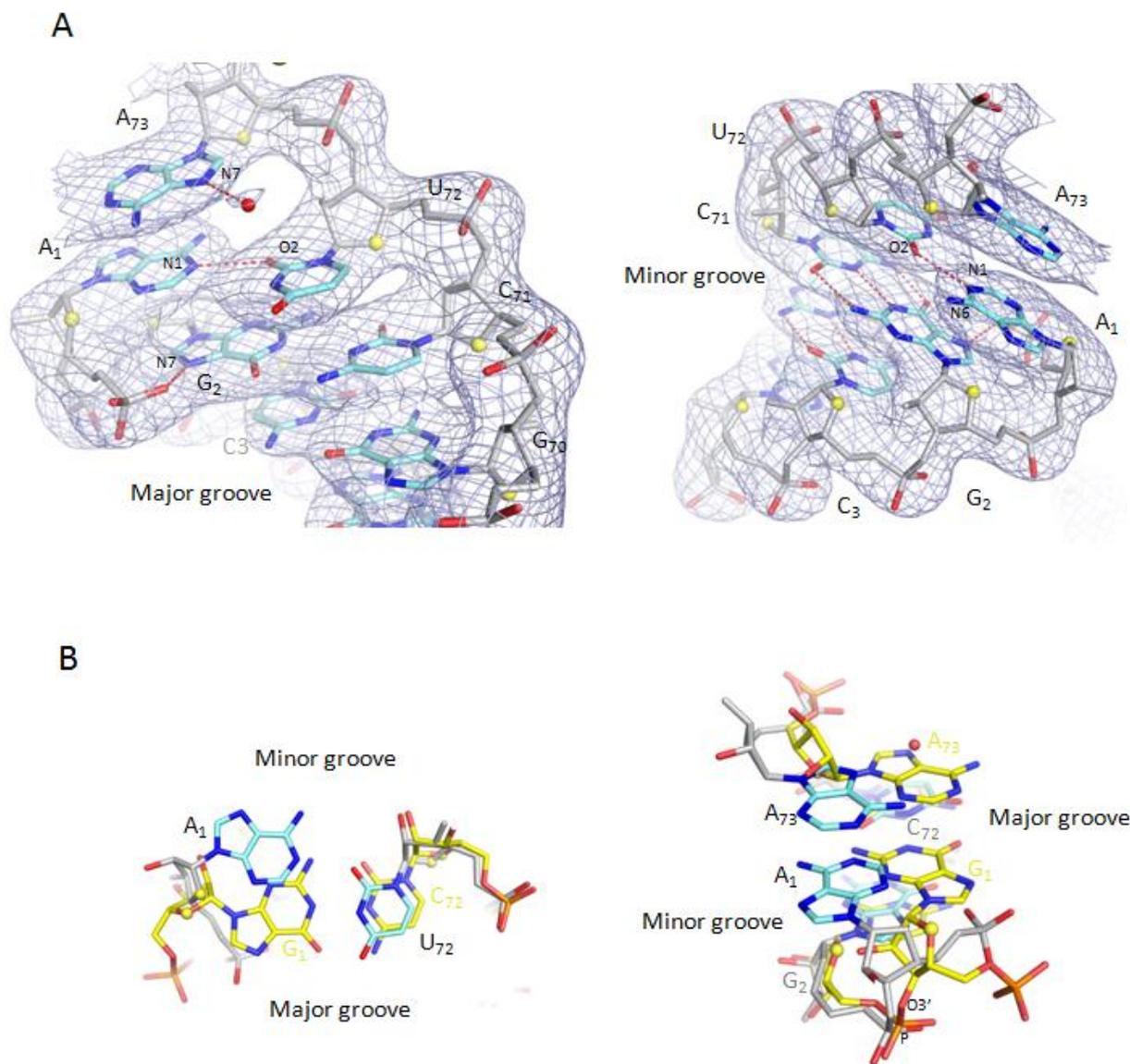


Figure 4

## 2.2 Discussion de l'étude de l'ARNt<sup>Met</sup>A1-U72

La délétion des gènes d'ARNt<sub>f1</sub><sup>Met</sup> et d'ARNt<sub>f2</sub><sup>Met</sup> par la technique Lambda Red chez *E. coli* nous a permis d'exprimer et purifier en grande quantité l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>, dans le but d'obtenir sa structure cristallographique ou encore de l'utiliser lors de la reconstitution de complexes de démarrage archées. En effet, l'expression d'ARNt initiateur d'*E. coli* modifié dans ses bases 1 et 72 permet de pallier l'absence des gènes d'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli*. Cette technique est donc aussi un outil nous permettant de tester la fonctionnalité d'un ARNt initiateur mutant chez *E. coli*. En effet, la croissance de la souche PalΔZWV est plus lente sur boîte de pétri que la souche non modifiée (MG1655). Nous pouvons donc utiliser la souche PalΔZWV pour tester la capacité d'ARNt initiateurs à pallier le défaut de croissance.

De manière intéressante, lorsque nous exprimons l'ARNt initiateur de l'archée *Sulfolobus solfataricus* dans la souche PalΔZWV, nous observons que la croissance rapide est rétablie (voire figure 2.2 c/). Cet ARNt peut donc être utilisé lors du démarrage de la traduction chez *E. coli*, même s'il ne semble pas provoquer pas un retour complet à la croissance optimale d'*E. coli*.

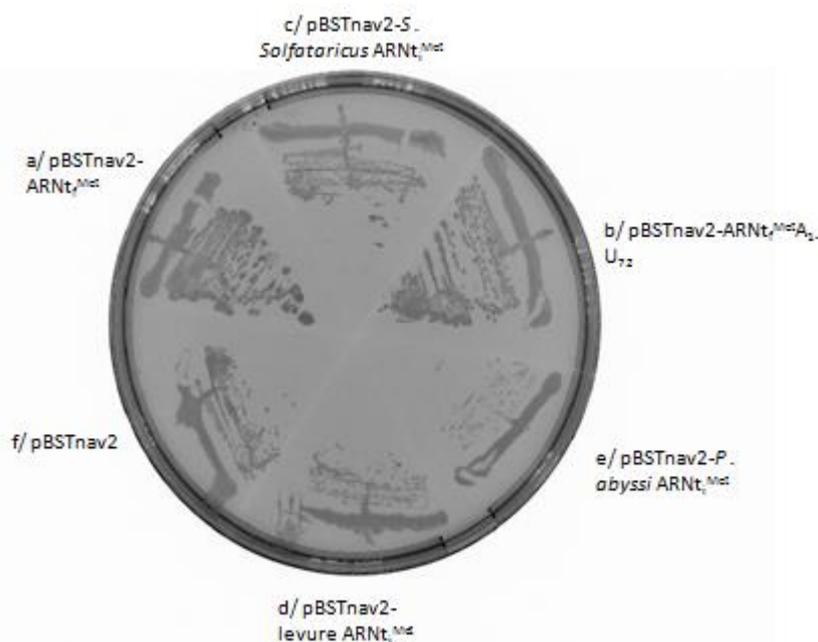


Figure 2.2 : tests de complémentation de la souche PalΔZWV avec différents ARNt.

Au contraire, lorsque cette souche PalΔZWV est transformée avec un plasmide portant le gène d'ARNt initiateur de levure ou de l'archée *P. abyssi*, on observe la même croissance sur boîte de pétri que lorsque cette souche est transformée avec un plasmide témoin ne contenant pas de gène d'ARNt<sub>i</sub> (voir figure 2.2 d/, e/ et f/). Ces résultats nous indiquent que les ARNt<sub>i</sub> de levure ou de *P.*

*abyssi* ne peuvent pas être utilisés efficacement par *E. coli* pour le démarrage de la traduction, ou bien qu'ils ne sont pas produits à un niveau suffisant. Nous avons toutefois vérifié que la présence du plasmide produisant l'ARNt de *P. abyssi* dans une souche standard conduisait bien à une surproduction des ARNt accepteur de méthionine d'un facteur environ 5.

L'ARNt<sub>i</sub> de levure possède une séquence nucléotidique plus éloignée de celle de l'ARNt<sub>f</sub><sup>Met</sup> d'*E. coli* que celle des ARNt initiateurs archéens. En particulier la séquence de la boucle T est très différentes et contient une modification de base connue pour participer au repliement de cet ARNt (modification m1A58)(Anderson et al. 1998) (voir figure 2.3).

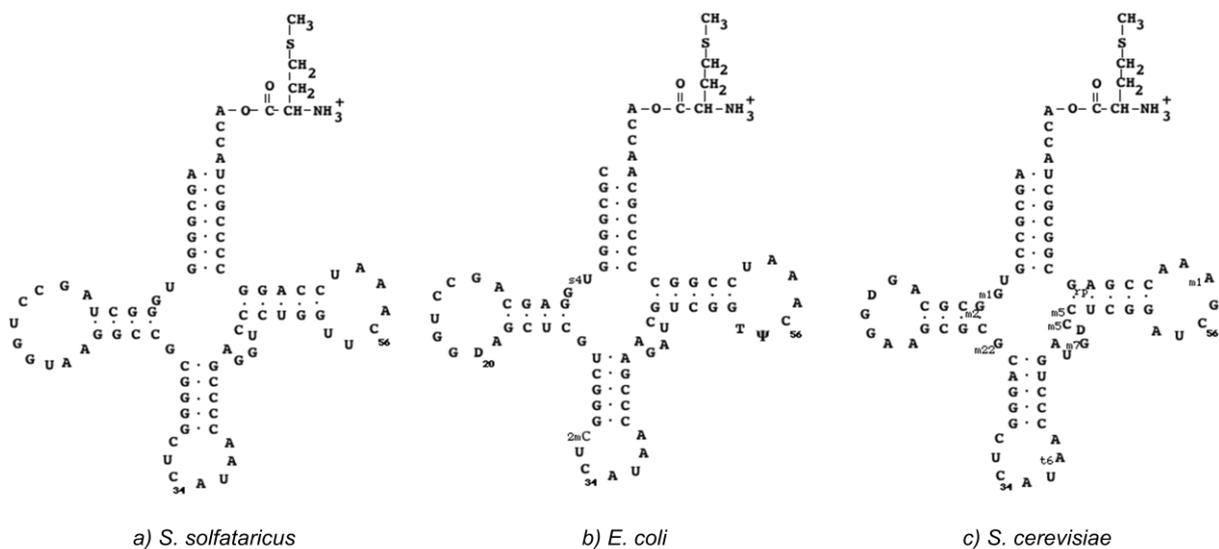


Figure 2.3 : Structure en feuille de trèfle d'ARNt initiateur archéen (a), bactérien (b) et eucaryote (c).

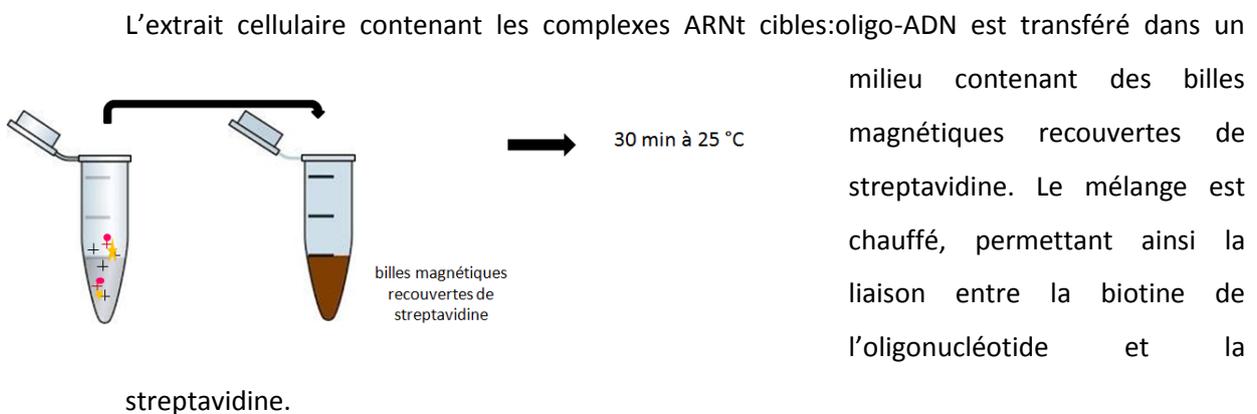
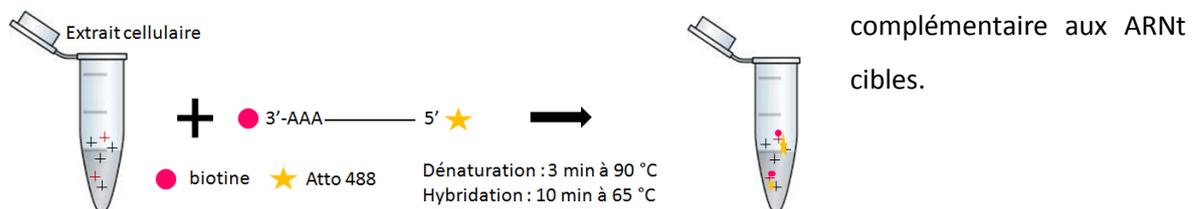
Une possibilité est que cet ARNt<sub>i</sub> de levure modifié chez *E. coli* ne puisse pas se replier correctement et donc être utilisé par la bactérie à la place de son ARNt initiateur naturel.

De la même façon, on peut imaginer que l'ARNt<sub>i</sub> de *P. abyssi* exprimé chez *E. coli* est mal replié. En effet, la température optimale de croissance de *P. abyssi* est de 96°C. Il est possible que le repliement de cet ARNt soit défavorable à 37°C. On peut aussi imaginer que l'ARNt<sub>i</sub> de *P. abyssi* produit chez *E. coli* soit hypomodifié et que l'absence d'une modification post-transcriptionnelle soit responsable de son mauvais repliement.

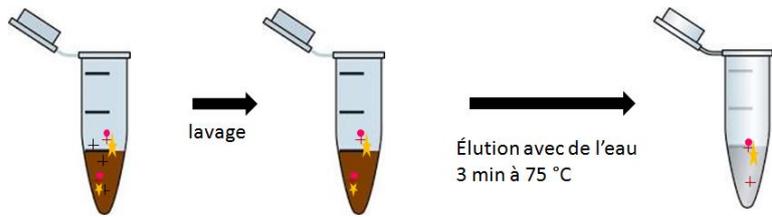
Dans ce contexte, nous avons commencé une collaboration avec le professeur Mark Helm de l'université de biologie Johannes Gutenberg-Universität à Mainz en Allemagne. Le professeur Mark Helm est un spécialiste des nucléosides modifiés dans les acides nucléiques et en particulier dans les ARNt. Jusqu'à présent, très peu de séquence d'ARNt d'archées sont connues. En particulier, les modifications de base des ARNt méthionine n'ont jamais été étudiées pour l'archée *Pyrococcus*

*abyssi*, ni pour l'archée *Sulfolobus solfataricus*. De nouvelles techniques telles que le séquençage dernière génération, la chromatographie et la spectrométrie de masse qui sont développées dans le laboratoire du Pr Marc Mark Helm ont ouvert de nouvelles perspectives pour l'étude des modifications biochimiques des ARNt. J'ai passé 6 semaines dans le laboratoire du Pr. Helm dans le but de purifier puis d'étudier par spectrométrie de masse les ARNt initiateur et élongateur méthionine de *P. abyssi* et de *S. solfataricus*. En effet, les cartes de modifications post-transcriptionnelles chez ces ARNt s'avèrent être inconnues et pourraient nous permettre d'identifier des modifications de bases pouvant expliquer comment ces deux ARNt sont différenciés par les systèmes de démarrage et d'allongement chez les archées. Dans ce cadre, j'ai pu extraire et purifier en quantité conséquente les ARNt initiateur et élongateur méthionine à partir des cellules de nos archées d'intérêt (voir Matériels et Méthodes).

Pour ce faire, les cellules provenant de culture des archées *P. abyssi* ou *S. solfataricus* ont été cassées et la fraction soluble contenant les ARNt a été récupérée. Des oligonucléotides d'ADN totalement complémentaires de l'ARNt cible et marqués en 3' avec de la biotine et en 5' avec de l'Atto488 (pour permettre leur détection sur gel ultérieurement) sont ajoutés dans l'extrait cellulaire. Ces oligonucléotides possèdent en plus de la séquence complémentaire de l'ARNt cible, 2 ou 3 adénines supplémentaires en 3' pour faciliter leur liaison aux billes magnétiques utilisées par la suite. Une étape de dénaturation et d'hybridation permet aux oligonucléotides de venir se lier de façon

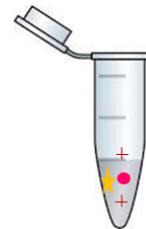


Le mélange contenant les billes est lavé afin de ne conserver que les complexes ARNt cibles:oligo-ADN, puis élué avec de l'eau après 3min à 75°C afin de s'affranchir des billes magnétiques.



L'extrait purifié d'ARNt cible, ARNt cible:oligo-ADN est alors mis en présence de DNase I afin de dégrader les oligos-ADN toujours présents dans le milieu, qui pourraient interférer avec des futures expériences.

Dnase I  
2h at 37 °C



La purification des ARNt cibles est vérifiée après migration sur un gel 10% dénaturant.

PAGE 10% dénaturant des extraits d'ARNt.



Scan 1 : gel non coloré → seul les oligos d'ADN marqués à l'Atto488 sont détectable (Contrôle de la digestion par la DNase)  
(excitation : 501 nm / emission : 526nm)

Scan 2 : gel coloré → excision des ARNt cible sur gel [excitation : 532 nm ou 610 nm / emission : (<526 nm)]

On peut observer sur le gel non coloré (à gauche) que les oligos d'ADN ne sont plus présents dans l'extrait contenant nos ARNt cibles purifiés. Le gel est alors coloré avec du GelRed, permettant ainsi de mettre en évidence la présence de nos ARNt cibles sur le gel, et ainsi pouvoir les extraire en excisant les bandes de gel les contenant. Pour finir, les ARNt cibles sont extraits du gel et conservés à -20°C après une précipitation à l'éthanol.

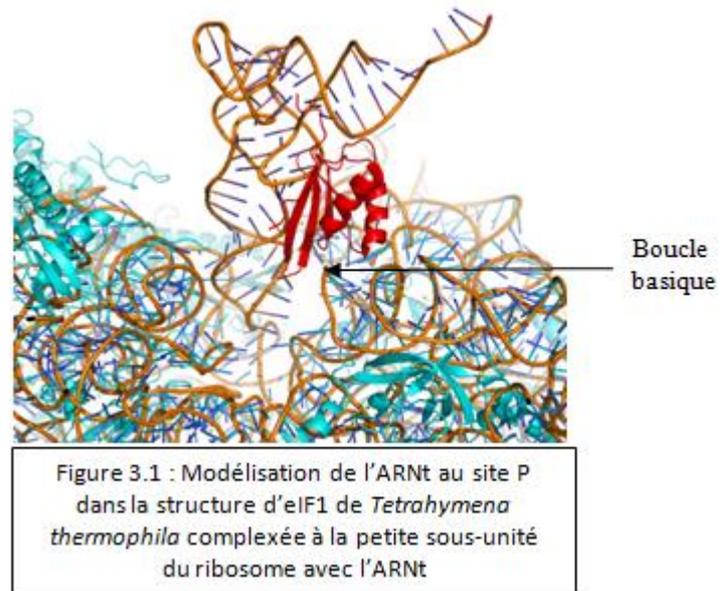
Par cette méthode, nous avons donc pu extraire et purifier les ARNt méthionines élongateur ou initiateur des archées *P. abyssi* ou *S. solfataricus*, nous permettant de poursuivre nos études.

Les premières études des ces ARNt par spectrométrie de masse nous ont permis de mettre en évidence des modifications de bases. Cependant, les études restent à approfondir afin de pouvoir caractériser ces modifications et de les localiser dans la séquence des ARNt.

# **3 Etude du rôle d'aIF1 dans le complexe de démarrage de la traduction**

Le facteur eIF1 joue un rôle central dans le démarrage de la traduction eucaryote. En effet, il favorise le recrutement du complexe ternaire eIF2-GTP-Met-ARNt<sub>i</sub><sup>Met</sup>, et sa présence dans le complexe de démarrage permet de prévenir le démarrage de la traduction sur un mauvais codon ou dans un mauvais contexte (Pestova et al., 1998). Chez les eucaryotes, les facteurs eIF1 et eIF1A ont été montrés nécessaires au « balayage » de l'ARNm, vraisemblablement parce que leur liaison induit une conformation particulière du ribosome, dite « ouverte », compétente pour le « balayage ». Cette fonction pourrait être modulée par l'interaction de ces protéines avec d'autres partenaires tels que eIF3 et eIF5. Chez les archées, il n'y a pas de « balayage » de l'ARNm sur une longue distance. En effet, la petite sous-unité du ribosome est pré-positionnée à proximité du codon de démarrage grâce à l'interaction SD:antiSD ou bien le démarrage de la traduction implique la liaison d'un complexe sur une région 5' non traduite très courte. De plus, les facteurs eIF3 et eIF5 sont absents. Ainsi, les fonctions moléculaires d'e/aIF1 et e/aIF1A auraient, dans une certaine mesure, divergées au cours de la différenciation des branches eucaryotes et archées du vivant.

Au début de mon travail de thèse, la détermination très récente de la structure d'eIF1 de *Tetrahymena thermophila* (Tt-eIF1) complexée à la petite sous-unité du ribosome (Rabl et al, 2011) suggérait que le site de liaison d'eIF1 sur le ribosome se situe juste devant le site de liaison de l'ARNt initiateur au site P. Grâce à cette structure on a pu modéliser la position de l'ARNt au site P dans le



complexe 40S:eIF1. Dans ce modèle, la position d'eIF1 paraît idéale pour permettre au facteur de vérifier la justesse du codon et de l'ARNt choisis pour démarrer. Une boucle d'eIF1 contenant des résidus basiques est en contact visible avec l'ARN ribosomal. Dans la structure cristallographique, ce contact implique seulement un résidu basique de cette boucle, alors que les autres résidus, moins bien définis dans la densité électronique, pourraient interagir avec l'ARNm ou avec l'ARNt (Rabl et al, 2011). Ces résidus pourraient interagir au niveau de la tige anticodon de l'ARNt, suggérant ainsi un mécanisme expliquant la discrimination des ARNt initiateurs et des ARNt élongateurs par ce facteur. En effet, les trois paires de bases GC caractéristiques des ARNt initiateurs se retrouvent dans la tige anticodon que ne possèdent pas les ARNt élongateurs.

Quand j'ai commencé mon travail, peu de données existaient sur le rôle de l'homologue archéen aIF1 dans le démarrage de la traduction archée. Deux études avaient permis de montrer que le facteur de démarrage aIF1, assisté par aIF1A, permettrait de stimuler la liaison du CT à la petite sous-unité du ribosome (Hasenörhl D et al., 2006 ; Hasenörhl D et al., 2009).

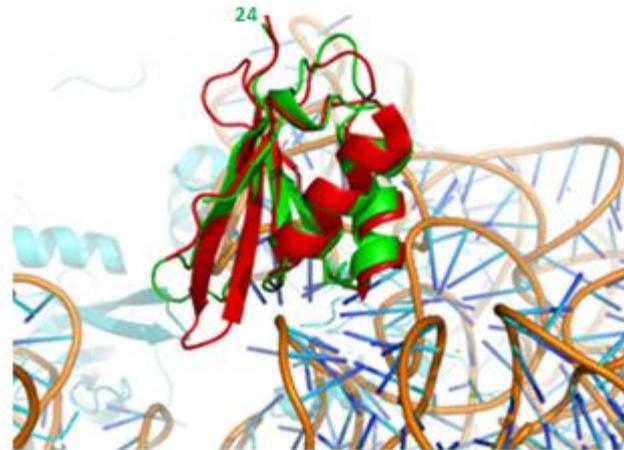


Figure 3.2 : Modélisation de la structure d' eIF1 (rouge) et aIF1 (vert) obtenue par cristallographie aux rayons X sur la sous-unité 40S du ribosome

Juste avant mon arrivée également, la

structure cristallographique du facteur archée aIF1 de *Methanocaldococcus jannaschii* (Mj-aIF1) avait été obtenue au laboratoire. Cette structure se superpose très bien à celle du facteur eIF1 de *T. thermophila* en complexe avec la sous-unité 40S.

De manière intéressante, la boucle basique adopte une conformation différente. Cette différence de conformation suggérait un ajustement possible de cette région lors de la liaison de l'ARNt au site P et/ou d'aIF2.

Les facteurs purifiés aIF2, aIF1, aIF1A ainsi que la sous-unité 30S et des ARNm modèles de l'archée *P. abyssi* (dont les ARNm présentent le plus souvent une séquence SD) peuvent être facilement obtenus au laboratoire. *Pyrococcus abyssi*, qui peut se développer jusqu'à 101.5°C, produit des protéines thermostables. Les protéines de *P. abyssi* produites par *E. coli* sont elles aussi thermostables. Cette propriété facilite grandement leur purification. De plus, des cultures à grande échelle (fermenteur de 160 litres) de cellules de *P. abyssi* avaient été réalisées avant mon arrivée. A partir de ces cellules, nous avons pu purifier des sous-unités ribosomales de cet organisme. Ainsi, la disponibilité de chaque composant nous permet de produire et d'isoler chaque intermédiaire du complexe de démarrage et d'envisager des études du rôle d'aIF1 dans la formation de ces complexes.

Mon travail a consisté à caractériser *in vitro* la fonction du facteur aIF1 dans le démarrage de la traduction archée. Pour cela, j'ai étudié dans un premier temps la capacité de liaison de ce facteur à la petite sous-unité du ribosome. Des variants d'aIF1, conçus à partir des données structurales ont été purifiés et la liaison de ces variants au ribosome a été étudiée par anisotropie de fluorescence. Dans un second temps, en utilisant la méthode d'empreinte du ribosome (toeprinting), j'ai analysé

les acides aminés d'aIF1 impliqués dans la stabilisation du complexe de démarrage et dans la discrimination des mauvais complexes.

La confrontation des résultats obtenus en utilisant l'anisotropie de fluorescence et en utilisant la méthode de « toeprinting » nous permet d'identifier les acides aminés de cette protéine impliqués dans la liaison au ribosome et ceux impliqués dans la sélection du bon complexe de démarrage.

### **3.1 Étude de la liaison du facteur aIF1 sur la sous-unité du ribosome par anisotropie de fluorescence**

L'étude de la liaison d'aIF1 et de ses variants au ribosome par anisotropie de fluorescence nécessite la reconstitution préalable d'un complexe de démarrage archée fonctionnel in vitro.

#### **3.1.1 Obtention des acteurs du complexe de démarrage et principe de l'anisotropie de fluorescence**

##### 3.1.1.1 Obtention de petites sous-unités ribosomales 30S

Pour cela, l'obtention de fractions très pures de sous-unité 30S ribosomales d'archée est nécessaire. La présence au laboratoire d'extraits cellulaires de l'archée *Pyrococcus abyssi* nous a permis de mettre au point un protocole permettant d'extraire et de purifier les sous-unités ribosomales de celle-ci.

La purification de ces sous-unités ribosomales repose sur une technique développée dans les années 1960 d'isolement de ribosome bactérien (Britten et Roberts 1960). Le ribosome bactérien, tout comme le ribosome archée par ailleurs, et les sous-unités ribosomales le constituant, ont des constantes de sédimentation d'environ 30S, 50S et 70S (S=Svedberg). La séparation de ces différents constituants d'un extrait cellulaire peut être obtenue par une sédimentation après ultracentrifugation sur un gradient de saccharose stabilisé. La sédimentation de ces constituants sur des gradients de saccharose linéaires permettent d'isoler les polysomes, le ribosome entier 70S, ainsi que les sous-unités 30S et 50S.

Dans un premier temps, les cellules de l'archée *P. abyssi* sont resuspendues dans un tampon à pH7.5 contenant 10mM d'acétate de magnésium et 500mM de NH<sub>4</sub>Cl (pour décrocher les facteurs liés au ribosome) auquel sont ajoutées des billes de verre pour le cassage des cellules. Les vibrations créées par un vibreur permettent le cassage de la paroi des cellules par les billes de verre. Une fois les cellules cassées, celles-ci sont centrifugées et le surnageant contenant les ribosomes est récupéré. Ce surnageant est réparti sur un coussin de saccharose et ultracentrifugé, permettant ainsi d'obtenir dans le culot des fractions de ribosome affranchies de protéines et d'ARN cellulaire. Ce culot contenant des ribosomes, mais aussi des polysomes, est resuspendu avec du tampon de resuspension puis appliqué sur une colonne de chromatographie de tamisage moléculaire. Cette étape permet de polir la préparation et d'éliminer en partie la phosphoenol pyruvate synthase (Mutchins et al., 2001), une enzyme multimérique de haut poids moléculaire (octamère de 8×92 kDa)

co-purifiant avec les sous-unités ribosomales. Les fractions contenant les ribosomes sont regroupées puis dialysées dans un tampon à pH7.5 contenant cette fois 2.5mM de magnésium. La diminution de cette concentration en magnésium permet la dissociation des sous-unités ribosomales. La solution dialysée contenant maintenant un mélange de sous-unités 30S et 50S est déposée sur un gradient de saccharose puis ultracentrifugée, afin de pouvoir séparer grâce à leur différence de vitesse de sédimentation les sous-unités ribosomales. Les sous-unités purifiées peuvent alors être récupérées dans des fractions lors du prélèvement des gradients, puis ultracentrifugées et finalement conservées dans le tampon de resuspension.

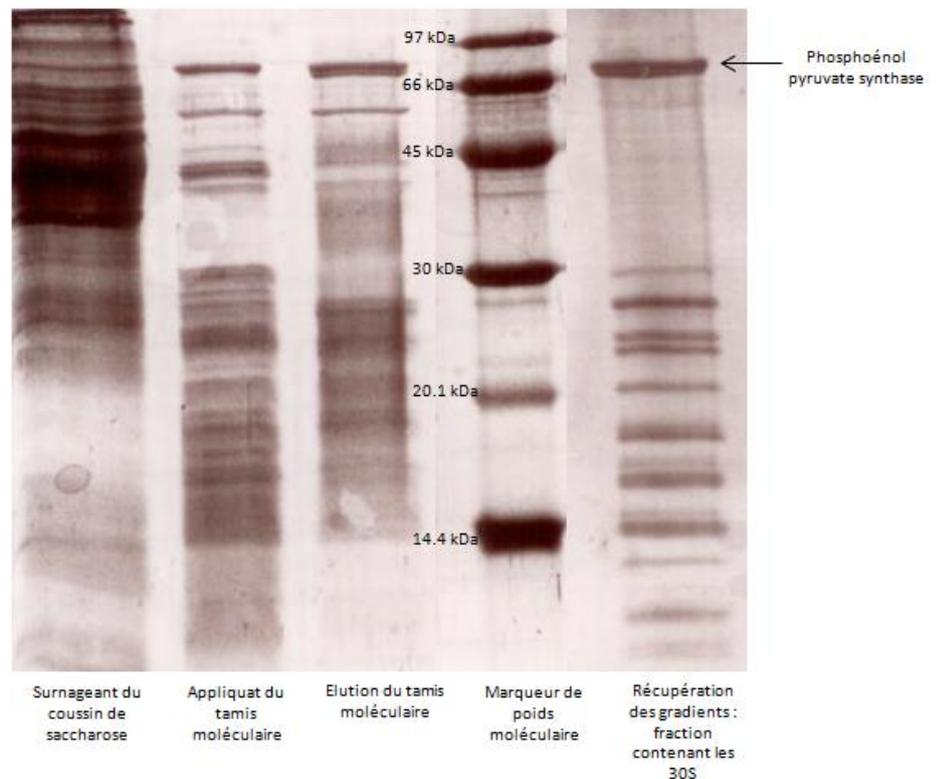


Figure 3.3 : Contrôle sur gel SDS-PAGE 16% des différentes étapes de purification des sous-unités ribosomales de *P. abyssi*

### 3.1.1.2 Purification du facteur de démarrage aIF1 et de ses variants

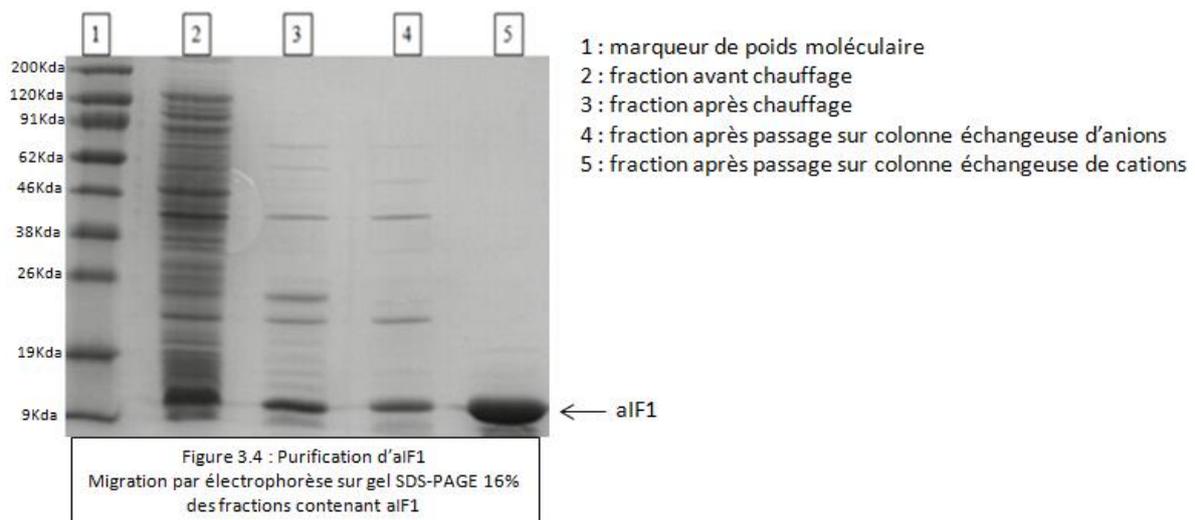
Afin de pouvoir étudier le rôle d'aIF1, il est nécessaire d'obtenir des fractions purifiées de ce facteur ainsi que des variants de celui-ci obtenus après mutagenèse dirigée du gène codant pour aIF1 par PCR (Polymerase Chain Reaction).

Pour cela, le facteur aIF1 ou ses variants sont exprimés dans une souche d'*E. coli*. Après désintégration aux ultrasons des cellules, l'extrait cellulaire centrifugé est chauffé à 70°C afin d'éliminer les protéines d'*E. coli* thermolabiles.

Après centrifugation, le surnageant contenant les protéines restées solubles après chauffage (dont la protéine d'intérêt aIF1 de *P. abyssi*) est appliqué sur une colonne échangeuse d'anion préalablement équilibrée en tampon A10%B (100mM NaCl). aIF1 et ses mutants ont un point isoélectrique acide (de l'ordre de 6.6), mais la protéine possède des segments localisés riches en résidus basiques. Ainsi, nous constatons qu'aIF1 ne s'accroche pas sur résine échangeuse d'anion. Par contre cette étape permet de débarrasser notre échantillon de la majorité des acides nucléiques qui eux interagissent

fortement avec la résine. Les fractions correspondant au pic non retenu sur la colonne sont rassemblées et réappliquées sur une colonne échangeuse de cations préalablement équilibrée en tampon A10%B. Finalement, aIF1 et ses variants sont retenus sur cette colonne vraisemblablement *via* l'interaction des segments basiques avec la résine. Les protéines sont éluées par un gradient linéaire de NaCl. Les fractions correspondantes sont collectées, rassemblées, concentrées à 100µM, puis stockées à -80°C.

Les variants d'aIF1 dans lesquels un résidu basique a été modifié ne sont plus retenus sur la résine échangeuse de cations. Les fractions correspondant au pic non retenu de la colonne échangeuse de cations sont alors rassemblées, concentrées dans 500µL, puis appliquées sur un tamis moléculaire de type Superdex 75 équilibrée en tampon A10%B.



### 3.1.2 L'anisotropie de fluorescence : principe général

Une première étude de la liaison d'aIF1 à la sous-unité 30S du ribosome peut maintenant être envisagée. Nous avons choisi d'utiliser la méthode de l'anisotropie de fluorescence. Pour cela nous avons dû implémenter la méthode au laboratoire.

L'anisotropie ( $r$ ) est le terme décrivant la mesure de la polarisation de rayons lumineux émis par des fluorophores par exemple. Ainsi, un échantillon présentant une anisotropie non nulle affiche une émission de rayons polarisés. Dans une solution homogène, les fluorophores à l'état fondamental sont orientés de manière aléatoire. Quand ils sont exposés à une lumière polarisée, ces fluorophores émettent des rayons polarisés dans la direction du rayon incident. On peut cependant observer une dépolarisation de l'émission lors de la diffusion rotationnelle du fluorophore. La mesure de

l'anisotropie permet de caractériser la moyenne de déplacement angulaire du fluorophore produit entre l'absorption et l'émission d'un photon.

Dans la plupart des expériences d'anisotropie, les échantillons sont excités par une lumière polarisée verticalement. L'intensité de l'émission de lumière est mesurée par un polarisateur. Quand l'émission polarisée est orientée parallèlement ( $\parallel$ ) à la direction de l'excitation polarisée, l'intensité observée est appelée  $I_{\parallel}$ . Au contraire, quand le polarisateur est perpendiculaire ( $\perp$ ) à l'excitation, l'intensité mesurée est appelée  $I_{\perp}$ .

L'anisotropie est définie comme suit :

$$r = (\text{intensité de fluorescence polarisée verticalement} - \text{intensité polarisée horizontalement}) \div (\text{intensité verticalement} + 2 \times \text{intensité polarisée horizontalement}) = (I_{\parallel} - I_{\perp}) / (I_{\parallel} + 2I_{\perp})$$

Le taux de diffusion rotationnelle dépend de la viscosité du solvant, mais aussi de la taille et de la forme de la molécule en rotation. Ainsi, pour des petits fluorophores couplés à une protéine en rotation libre dans un milieu peu visqueux, le taux de diffusion rotationnelle est plus rapide que le taux d'émission. Dans ces conditions, l'émission est dépolarisée et l'anisotropie est proche de 0 (voir figure 3.5)

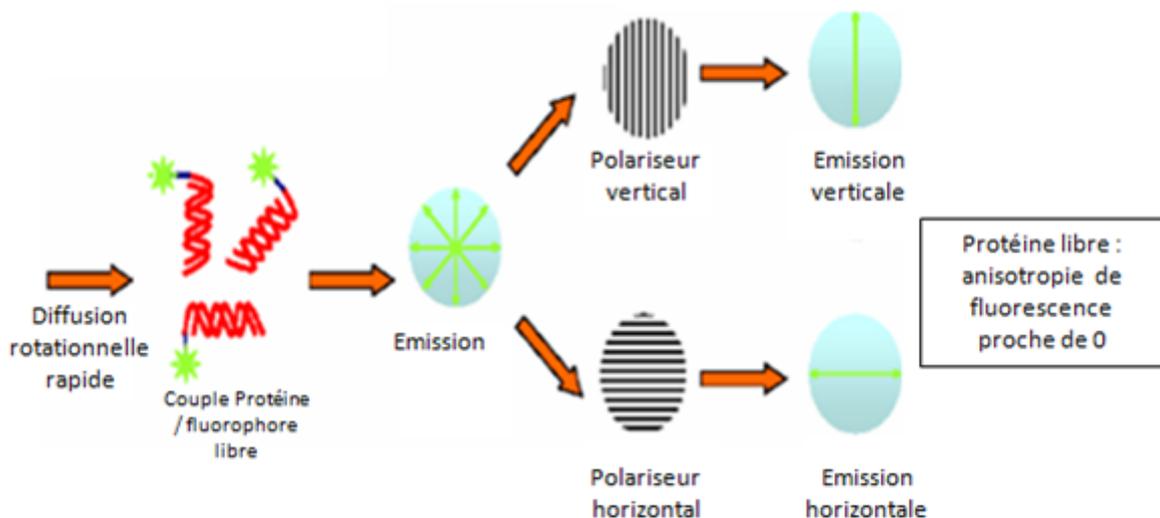


Figure 3.5 : anisotropie de fluorescence proche de 0

Au contraire, lorsque la vitesse de diffusion rotationnelle diminue, l'émission de lumière par le fluorophore présente une répartition privilégiée, le taux d'émission verticale et horizontale est alors différent et l'anisotropie augmente (voir figure 3.6)

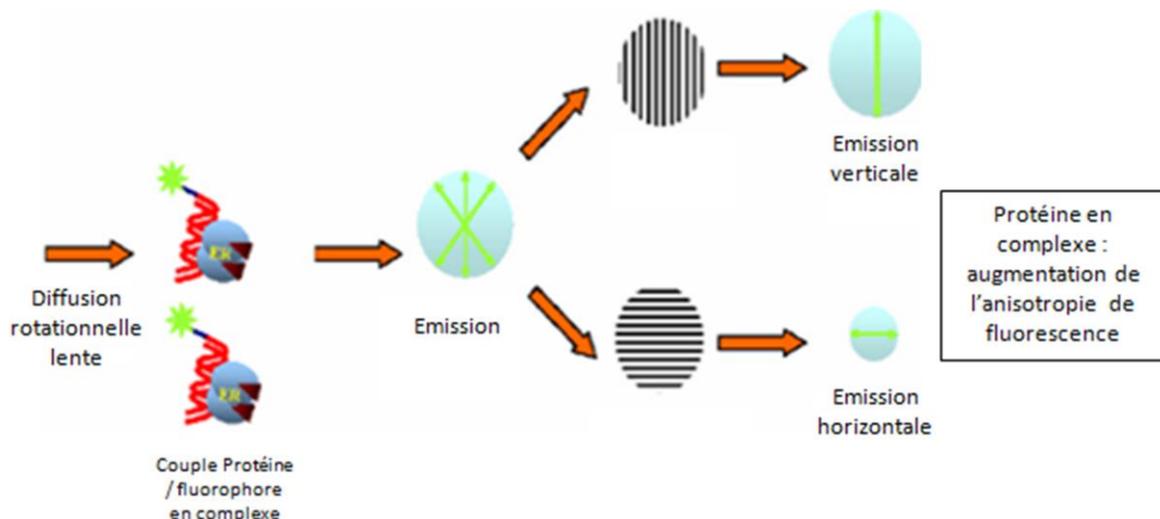


Figure 3.6: anisotropie de fluorescence  $\neq 0$

La variation de l'anisotropie de fluorescence lors du déplacement d'un fluorophore dans un milieu donné a permis de nombreuses applications de cette technique en biochimie. Ainsi, nous avons utilisé l'anisotropie de fluorescence pour observer la liaison d'aIF1 ou de ses variants à la petite sous-unité du ribosome.

Comme nous avons pu le voir plus haut, la liaison d'une protéine à une autre molécule entrainera une variation de ses propriétés physiques en faisant varier sa rotation jusqu'alors libre dans le milieu, augmentant ainsi son anisotropie. La mesure de la variation de l'anisotropie de fluorescence d'une protéine couplée à un fluorophore dans un milieu donné est donc un moyen indirect d'observer la liaison de cette protéine à une molécule connue. Une protéine libre marquée avec un fluorophore émet de la fluorescence de manière plus isotrope que lorsqu'elle est complexée avec d'autres molécules. Ainsi, l'anisotropie de fluorescence est plus élevée si la protéine est en complexe avec une autre molécule. Dans cette étude, si un complexe se forme, le facteur aIF1 de *Pyrococcus abyssi* (108 acides aminés, Mm=11,228kDa) ou l'un de ses variants sera immobilisé sur la petite sous-unité du ribosome (Mm=694kDa). On s'attend donc à une grande variation de l'anisotropie si le complexe est formé.

### 3.1.2.1 Marquage des protéines

L'utilisation de cette technique nécessite le marquage de notre protéine d'intérêt avec un fluorophore. Un dérivé maléimide de la coumarine est utilisé : le Mdcc ou 7-Diethylamino-3-(((2-Maleimidyl)ethyl)amino)carbonyl)coumarine. La réaction de marquage des protéines avec le Mdcc correspond à la formation d'une liaison covalente entre le groupement maléimide du Mdcc et le

groupement soufre d'un résidu cystéine d'aIF1. La fraction de protéine marquée est appliquée sur une colonne échangeuse de cation ou un tamis moléculaire selon le mutant d'aIF1 étudié. Cette purification permet d'obtenir des fractions contenant exclusivement la protéine d'intérêt marquée (mélangée à la fraction non marquée) et d'exclure le Mdcc resté libre.

### 3.1.3 Mise au point du titrage d'aIF1 par la petite sous-unité 30S du ribosome

#### 3.1.3.1 Titration de l'ARNm modèle avec la sous-unité 30S du ribosome

Nous allons étudier par anisotropie de fluorescence la formation d'un complexe entre le facteur aIF1 et la sous-unité ribosomale préalablement complexée à un ARNm modèle de 26 bases correspondant à la région de démarrage de la traduction de l'ARNm codant le facteur aEF1A de *P. abyssi*. Cet ARNm contient une séquence SD forte (susceptible de s'apparier à 9 bases de l'ARNr 16S)(voir figure 3.7). Cet ARNm modèle a été synthétisé chimiquement (Dharmacon GE Healthcare).



figure 3.7 : Séquence de l'ARNm modèle utilisé pour les études par anisotropie de fluorescence

Dans un premier temps, la liaison de l'ARNm modèle à la sous-unité 30S doit être vérifiée, afin de pouvoir utiliser ce complexe dans des conditions optimales de liaison pour la suite de nos expériences. Pour cela, l'ARNm modèle est marqué avec de la fluorescéine (noté ARNm-fluo) (excitation de la fluorescéine à 485nm ; maximum d'émission de la fluorescéine à 520nm) à l'extrémité 3' afin de ne pas inhiber la liaison ARNm:30S. Une solution de 220µL de tampon de titrage à la concentration finale de 50nM d'ARNm-fluo a été préparée. Ce tampon de titrage contient du magnésium nécessaire à la formation des complexes ribosomaux et du tween 20 permettant d'éliminer l'adsorption des protéines et ARN sur la cuve en quartz. L'anisotropie initiale de cette solution est de 0.0269. La liaison de l'ARNm-fluo à la 30S est testée par des ajouts successifs de 1.7µL d'une solution à 672nM en 30S dans le tampon de titrage. L'expérience est réalisée à 51°C. Une première phase d'augmentation linéaire de l'anisotropie en fonction de la concentration en 30S dans le milieu est observée, puis la courbe adopte une tendance hyperbolique jusqu'à atteindre une phase stationnaire (voir figure 3.8). L'allure générale de cette courbe suggère une constante de dissociation (Kd) très faible au regard de la concentration en ARNm-fluo utilisé (Kd<<50nM). Les données ont

malgré tout été traitées avec le logiciel OriginPro dans lequel nous avons implémenté l'équation décrite un peu plus loin pour le titrage d'aIF1. Le traitement indique un  $K_d$  de  $3.6\text{nM} \pm 0.6\text{ nM}$ . Ces données nous permettent de poursuivre nos expériences d'anisotropie dans les conditions optimales de la formation d'un complexe ARNm:30S à  $51^\circ\text{C}$ , et de s'assurer par la suite une saturation de la sous-unité 30S en ARNm.

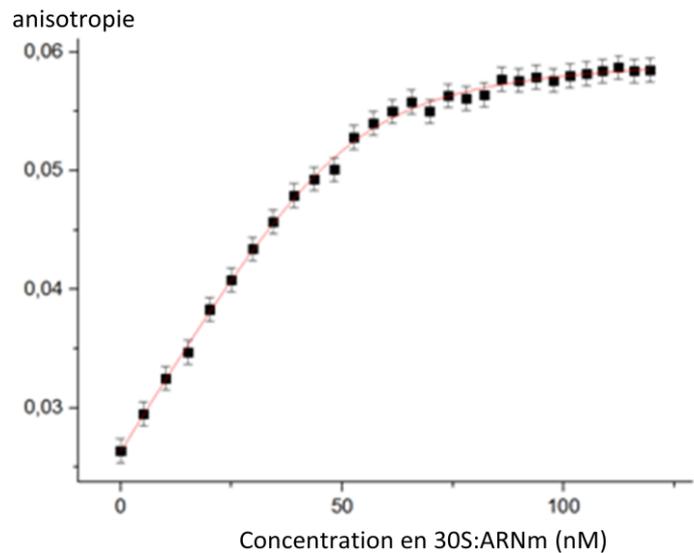


Figure 3.8 : Titrage de l'ARNm-fluo (50nM) avec 30S. Chaque mesure correspond à un ajout de 5nM de 30S

Par ailleurs, pour s'assurer que la variation d'anisotropie observée était bien due à la liaison spécifique de l'ARNm à la 30S, nous avons vérifié que l'ARNm non marqué était compétiteur de l'ARNm-fluo. Pour cela, des ajouts successifs de 10nM d'ARNm sont effectués en fin de titrage. Une diminution de l'anisotropie est observée jusqu'à un retour à la valeur d'anisotropie d'aIF1\* seul, permettant de vérifier la liaison spécifique de l'ARNm à la petite sous-unité 30S dans nos conditions expérimentale.

Finalement, une solution contenant l'ARNm-fluo (50nM) à également été titrée par du tampon de titrage afin de vérifier que la variation d'anisotropie observée n'était pas due à la dilution.

### 3.1.3.2 Titrage du facteur aIF1 de *P. abyssi* marqué sur la cystéine 63 avec la petite sous-unité 30S du ribosome et l'ARNm

Afin d'étudier la formation d'un complexe entre la petite sous-unité du ribosome et le facteur de démarrage aIF1 par anisotropie de fluorescence, ce dernier doit être marqué par un fluorophore, ici le Mdcc, se liant sur un résidu cystéine. Le facteur aIF1 possède une unique cystéine en position 63 (voir figure 3.9). D'après la structure du complexe 40S : eIF1 de *T. thermophila*, nous avons suspecté que cette cystéine 63 pouvait se trouver très proche de la petite sous-unité ribosomale et qu'elle pourrait donc être impliquée dans la formation du complexe aIF1:30S (voir figure 3.2 plus haut).

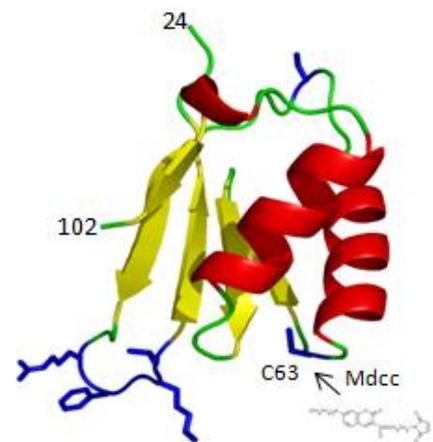


Figure 3.9 : Structure du facteur aIF1 de l'archée *M. jannaschii*

Ainsi, l'introduction du fluorophore très encombrant à cette position est susceptible d'interférer avec

la formation du complexe 30S : aIF1. Malgré tout, un premier test a été réalisé en utilisant de la protéine sauvage marquée.

La protéine marquée en position 63 est donc utilisée pour une première expérience de titrage. Pour cela, une solution de 320µL de tampon titrage (voir matériels et méthodes) à la concentration finale de 0.2µM d'aIF1 marquée a été préparée. L'anisotropie de cette solution est de 0.144 (excitation du Mdcc à 428nm ; maximum d'émission du Mdcc à 473nm). La liaison d'aIF1 au ribosome a ensuite été testée par des ajouts successifs d'un mélange contenant la sous-unité 30S de *P. abyssi* et l'ARNm. La solution de 30S:ARNm a été préparée à la concentration finale de 672nM en 30S, 1340nM d'ARNm, et le mélange 30S:ARNm est chauffé 2 minutes à 51°C pour favoriser la liaison de l'ARNm.

Le résultat de l'expérience est présenté dans la figure 3.10 ci-contre. Malgré l'atteinte d'une concentration de 161nM en 30S, aucune augmentation significative de l'anisotropie ne peut être observée.

Différentes hypothèses peuvent alors être proposées :

- (i) La méthode n'est pas adaptée à l'étude de la formation du complexe 30S : aIF1 : ARNm
- (ii) La liaison d'aIF1 nécessite la présence d'un autre composant du démarrage de la traduction
- (iii) Le groupement fluorophore interfère avec la liaison

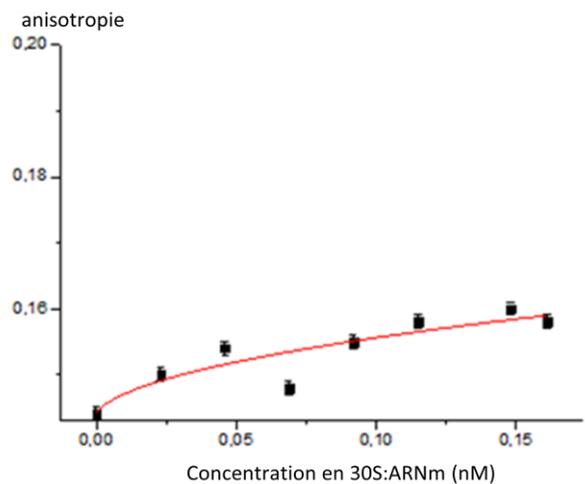


Figure 3.10 : Titrage d'aIF1WT (50nM) avec le complexe 30S:ARNm

Compte tenu du mode de liaison d'aIF1 à la sous-unité 30S, tel qu'il peut être suggéré à partir de la structure cristallographique d'eIF1:40S (figure 1), la troisième hypothèse semble la plus probable. Le marquage du facteur en position 63 ne serait donc pas approprié pour l'observation de la liaison d'aIF1 à la 30S. Nous avons donc cherché à construire un aIF1 mutant, possédant une cystéine unique à une position *a priori* plus favorable.

### 3.1.3.3 Titration du facteur aIF1\* marqué sur la cystéine 46 et mise au point de la mesure de la constante de dissociation du complexe 30S:ARNm:aIF1 par anisotropie de fluorescence.

aIF1 contient une boucle dans sa structure qui n'est visiblement pas en interaction directe avec le ribosome (voir figure ci-contre, boucle 2, lysine 46). La liaison d'un fluorophore dans cette boucle pourrait donc ne pas perturber l'interaction aIF1/30S. Cette boucle dont la séquence est très variable chez les aIF1 archées ne contient pas de cystéine permettant la liaison du Mdcc au facteur. Nous avons choisi de modifier la lysine en position 46 en cystéine. Cependant, afin de ne marquer aIF1 qu'en position 46, la cystéine 63 doit elle aussi être mutée afin que le Mdcc ne puisse plus se lier à ce site. Ainsi, nous avons choisi de muter cette cystéine en sérine. En effet, cette mutation est semi-conservative et une sérine en position 63 est rencontrée dans certaines séquences d'aIF1.

Une double mutation du gène codant pour aIF1 de *P. abyssi* a donc été effectuée. Une mutagenèse dirigée par PCR est réalisée avec un plasmide dans lequel est cloné le gène codant pour aIF1 et deux couples d'amorces. Les amorces permettent les mutations lysine 46 → cystéine (KC<sub>46</sub>) et cystéine 63 → sérine (CS<sub>63</sub>) (voir figure 3.11). J'ai vérifié la séquence des ADN plasmidiques des clones isolés après la PCR de mutagenèse afin de sélectionner les plasmides présentant les deux mutations. Le plasmide mutant est ensuite utilisé pour exprimer la protéine. Ce premier mutant est appelé aIF1\*.

Après expression d'aIF1\* dans *E. coli* BL21-rosetta, le facteur est purifié puis le marquage à la coumarine est réalisé (voir matériels et méthodes) avec un rendement de 54%. Ce rendement de marquage permet d'obtenir une intensité de fluorescence suffisante à une concentration de 50nM d'aIF1\* et ainsi de réaliser des mesures précises de l'anisotropie. Le titrage d'aIF1\* est réalisé par des ajouts successifs de 1.7µL d'une solution à 672nM de 30S et 1340nM d'ARNm (figure 3.12 ci-dessous).

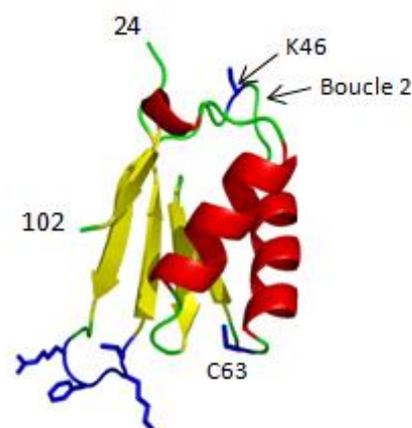


Figure 3.11 : Structure du facteur aIF1 de l'archée *M. jannaschii*

Une nette augmentation de l'anisotropie en fonction de la concentration en 30S dans le milieu est constatée. La courbe obtenue présente une première phase d'accroissement linéaire, puis un plateau de l'anisotropie lorsque la concentration de 30S:ARNm atteint 65nM (voir figure). Cette courbe nous suggère qu'aIF1\* se lie au complexe 30S:ARNm avec une constante de dissociation (Kd) très petite au regard de la concentration d'aIF1\* utilisée (50nM).

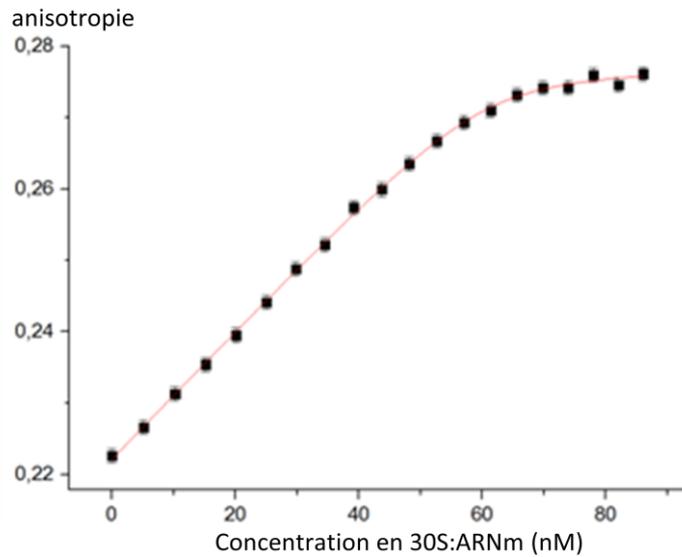


Figure 3.12 : Titrage d'aIF1\* (50nM) avec 30S:ARNm

La formation du complexe aIF1:30S suit la relation :



$$\text{avec } Kd = \frac{[aIF1]_{\text{libre}} [30S:ARNm]_{\text{libre}}}{[aIF1:[30S:ARNm]]} \quad (1)$$

D'autre part, l'anisotropie mesurée est reliée à la concentration de chaque espèce présente dans la solution par la relation :

$$r_{\text{mesurée}}[aIF1]_t = r_{aIF1}[aIF1] + r_{[aIF1:[30S:ARNm]]}[aIF1:[30S:ARNm]] \quad (2)$$

Cette relation est valable car on peut constater que l'intensité totale de fluorescence d'aIF1 ne varie pas au cours du titrage.

Lorsque le Kd est très petit devant la concentration d'aIF1 et que l'on est loin de la saturation, alors  $Kd \div [aIF1]_{\text{libre}} \ll 1$  et  $[30S:ARNm]_{\text{libre}} \approx 0$

Au cours du titrage, pratiquement chaque molécule de 30S:ARNm ajoutée dans la solution est liée à aIF1. L'anisotropie varie donc linéairement. Cette variation linéaire est stoppée brutalement lorsque la concentration en 30S:ARNm ajoutée devient égale à la concentration d'aIF1 totale.

Ainsi, l'analyse de cette courbe permet de déduire que la constante de dissociation ( $K_d$ ) d'aIF1\* pour la 30S est très inférieure à 50nM. Cette valeur ne peut pas être mesurée précisément. Pour pouvoir obtenir une détermination fiable de la constante de dissociation, nous devons diminuer la concentration d'aIF1 d'au moins un facteur 10 mais dans ce cas, nous n'avons malheureusement plus de signal exploitable.

Cependant, afin d'estimer le  $K_d$  d'aIF1\* au complexe 30S:ARNm, la courbe obtenue peut être analysée avec le logiciel d'analyse de données OriginPro, en utilisant les relations (1) et (2) précédente, et en considérant que la valeur de la constante de dissociation  $K_d$  est proche de la concentration d'enzyme, en écrivant la conservation des espèces, on obtient :

$$[aIF1]_{\text{totale}} = [aIF1]_{\text{libre}} + [aIF1:[30S:ARNm]]$$

$$[30S:ARNm]_{\text{totale}} = [30S:ARNm]_{\text{libre}} + [aIF1:[30S:ARNm]]$$

Pour simplifier, on appelle C la concentration en complexe aIF1:[30S:ARNm] et S la concentration en [30S:ARNm].

L'expression de l'anisotropie en fonction du  $K_d$  devient alors :

$$r[aIF1]_{\text{totale}} = r_0[aIF1]_{\text{totale}} + [r_{\infty}-r_0] \left( \frac{([aIF1]_T + S_T + K_d - \sqrt{([aIF1]_T + S_T + K_d)^2 - 4[aIF1]_T S_T})}{2} \right), \text{ où } r_{\infty} \text{ et } r_0 \text{ représente l'anisotropie de fluorescence d'aIF1 respectivement lié et libre.}$$

Cette formule appliquée à notre courbe conduit à un  $K_d$  entre aIF1 avec le complexe 30S:ARNm de  $0.9 \pm 0.3nM$ . Cette expérience a été réalisée de nouveau dans des conditions équivalentes et nous avons obtenu une courbe similaire dont le traitement par le logiciel OriginPro nous a permis d'estimer un  $K_d$  de  $0.53 \pm 0.24nM$  cohérent avec la valeur de  $K_d$  précédemment obtenue.

Cette valeur doit toutefois être considérée avec précaution au vu du  $K_d$  estimé et de la concentration en enzyme utilisée.

Dans le premier titrage, la variation est linéaire jusqu'à 65nM de complexe 30S:ARNm ajouté. Nous avons investi 50nM d'aIF1 dans le mélange. La valeur mesurée pour l'atteinte de la stœchiométrie est cohérente avec l'estimation théorique de la concentration d'aIF1 (estimée à partir d'un coefficient d'absorption molaire théorique très peu précis car la protéine ne possède pas de résidus tryptophane, voir matériels et méthodes). Cette valeur indique bien la formation d'un complexe 1:1

entre aIF1 et la sous-unité 30S. Elle est en faveur de l'établissement d'un complexe spécifique entre aIF1 et la 30S:ARNm.

Afin de s'assurer que la liaison observée était bien spécifique et n'était pas seulement liée à la présence du fluorophore, nous avons vérifié qu'aIF1\* non marqué était bien compétiteur de la version marquée. Pour ce faire, nous avons effectué un titrage retour en réalisant des ajouts successifs d'aIF1\* jusqu'à la concentration finale de 380nM. Nous avons pu constater un retour de l'anisotropie de fluorescence à une valeur de 0.226 qui correspond à l'anisotropie d'aIF1\* marqué libre (figure 3.13 ci-contre). Ceci montre qu'aIF1\* marqué a bien été chassé de son site de liaison sur la sous-unité 30S.

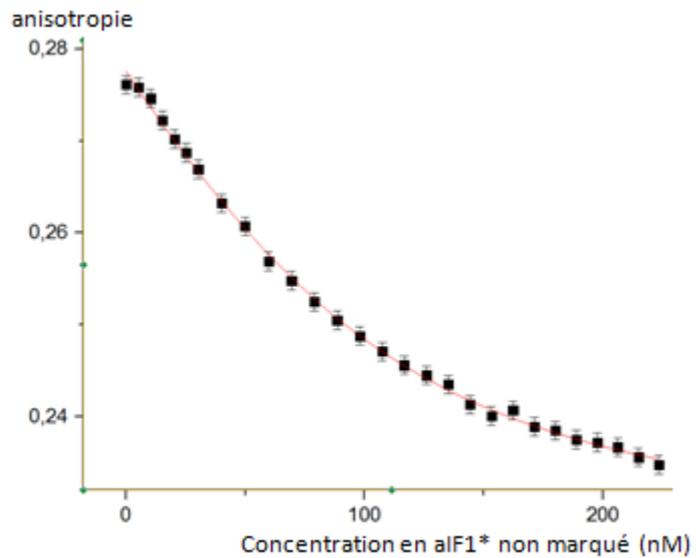


Figure 3.13 : Titrage retour d'aIF1\* complexé à la 30S:ARNm (50nM) avec aIF1\* non marqué

Ces résultats valident l'hypothèse selon laquelle c'était le marquage de la protéine aIF1 de *P. abyssi* sur la cystéine 63 qui perturbait la liaison du facteur à la sous-unité 30S. La construction d'un mutant de marquage (aIF1\*) permettant la liaison du fluorophore à aIF1 dans la boucle 2 permet d'observer la formation du complexe aIF1:30S par anisotropie de fluorescence. **La courbe obtenue montre que le Kd d'aIF1\* pour le complexe 30S:ARNm est très bas (<<50nM).**

Même si nous ne pouvons pas mesurer le Kd avec précision, cette méthode pourra être utilisée pour vérifier que les mutants d'aIF1 que nous étudierons ont gardé la capacité à se lier au ribosome. Pour cela, les mutations des acides aminés dont nous souhaitons tester l'importance seront réalisées dans le gène codant pour aIF1\*. Le fluorophore se liera ainsi toujours à la cystéine 46 de ces différents mutants.

### 3.1.3.4 Etude de la liaison d'aIF1\* à la sous-unité 30S seule.

La boucle de résidus basique du facteur aIF1, de par sa position dans la structure d'aIF1 en complexe avec la petite sous-unité peut interagir avec le ribosome, l'ARNt initiateur, mais aussi l'ARNm. Afin de tester l'implication de l'ARNm dans la liaison d'aIF1 à la 30S et ainsi déterminer l'interaction d'aIF1 avec l'ARNm, un titrage du facteur aIF1\* avec la sous-unité

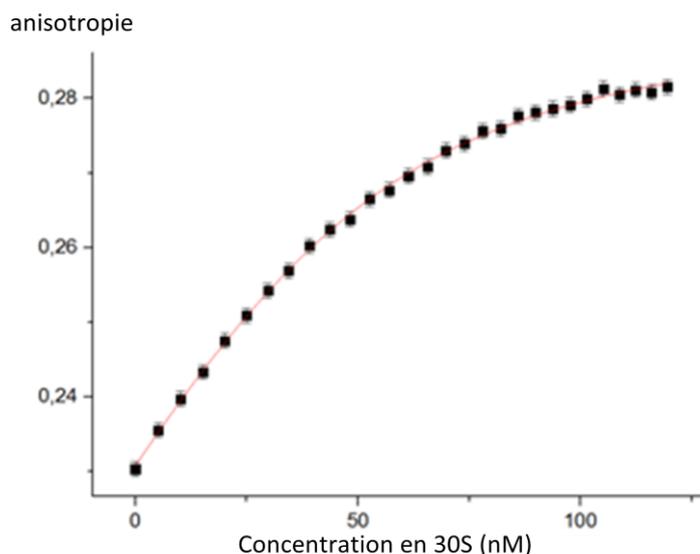


Figure 3.14 : Titrage d'aIF1\* (50nM) avec 30S

30S seule a été réalisé. Ce titrage est réalisé dans les mêmes conditions que précédemment sans que la 30S ne soit mise en complexe avec l'ARNm pour le titrage. La mesure est réalisée à 51°C avec une concentration de 50nM d'aIF1\*.

Cette fois, la courbe obtenue a une allure hyperbolique (voir figure 3.14). En utilisant les relations (1) et (2), lorsque la valeur de la constante de dissociation  $K_d$  est proche de la concentration d'enzyme, en écrivant la conservation des espèces, on obtient :

$$[aIF1]_{totale} = [aIF1]_{libre} + [aIF1:30S:ARNm]$$

$$[30S:ARNm]_{totale} = [30S:ARNm]_{libre} + [aIF1:30S:ARNm]$$

Pour simplifier, on appelle C la concentration en complexe aIF1:30S:ARNm et S la concentration en 30S:ARNm.

L'expression de l'anisotropie en fonction du  $K_d$  devient alors :

$$r[aIF1]_{totale} = r_0[aIF1]_{totale} + [r_{\infty}-r_0] \left( \frac{[aIF1]_T + S_T + K_d - \sqrt{([aIF1]_T + S_T + K_d)^2 - 4[aIF1]_T S_T}}{2} \right), \text{ où } r_{\infty} \text{ et } r_0 \text{ représente l'anisotropie de fluorescence d'aIF1 respectivement lié et libre.}$$

Ainsi, nous déterminons le  $K_d$  entre aIF1\* et la sous-unité 30S seule à 16 +/- 2nM.

La reproduction de cette expérience dans les mêmes conditions permet d'obtenir une courbe similaire et après traitement des données, un  $K_d$  équivalent de 12 +/- 4.9nM. Ces deux  $K_d$  sont

obtenus lorsque le nombre de site de fixation d'aIF1 sur la 30S n'est pas fixé à 1 site dans OriginPro, mais est défini par OriginPro. Le logiciel définit alors un nombre de site à  $1.23 \pm 0.08$ . Si le traitement de la courbe obtenue se fait cette fois avec un nombre de site défini et fixé à 1 site par 30S, alors le  $K_d$  augmente à  $25 \pm 4$ nM.

Cependant, en fixant ou non le nombre de site dans Origin lors du traitement de données, les résultats obtenus montrent que la présence de l'ARNm favorise la liaison d'aIF1 à la 30S. **Ceci suggère qu'aIF1 interagit avec la 30S et aussi avec l'ARNm.**

Compte-tenu de la position de la boucle basique d'eIF1 dans la structure du complexe 40S:eIF1, il est tentant d'imaginer que cette boucle basique soit impliquée dans la liaison de l'ARNm.

Afin de s'assurer de la spécificité de la liaison observée, une chasse d'aIF1\* par aIF1 non marqué est réalisée et on peut observer un retour à une valeur d'anisotropie se rapprochant de la valeur d'anisotropie d'aIF1\* libre dans le milieu.

Pour la suite des expériences, et afin de nous placer dans un contexte le plus favorable possible à la liaison étudiée, nous réaliserons les titrages d'aIF1 et de ses variants avec la sous-unité 30S préalablement complexée à l'ARNm.

#### **3.1.4 Caractérisation des résidus d'aIF1 impliqués dans la liaison au complexe 30S:ARNm**

Tous les mutants étudiés par la suite sont obtenus par mutagenèse dirigée par PCR du gène codant pour aIF1\*. Après vérification de la séquence complète du gène muté, la protéine d'intérêt est exprimée, purifiée par chromatographie sur colonne, puis marquée en utilisant le Mdcc (voir matériels et méthodes). Le volume initial du mélange aIF1/tampon tween (voir matériels et méthodes) est de 220 $\mu$ L. Chaque titrage est réalisé par ajout successifs de 1.7 $\mu$ L d'une solution contenant 672nM de 30S et 1340nM d'ARNm.

##### **3.1.4.1 Rôle de la boucle 30-35 dans la liaison d'aIF1 au complexe 30S:ARNm**

###### **3.1.4.1.1 Caractérisation du mutant aIF1\*-polyAla**

La boucle mobile d'aIF1 comprenant les résidus 30 à 35 semble être un bon candidat pour participer à la reconnaissance des ARNt initiateurs et de l'ARNm (voir figure 3.15).

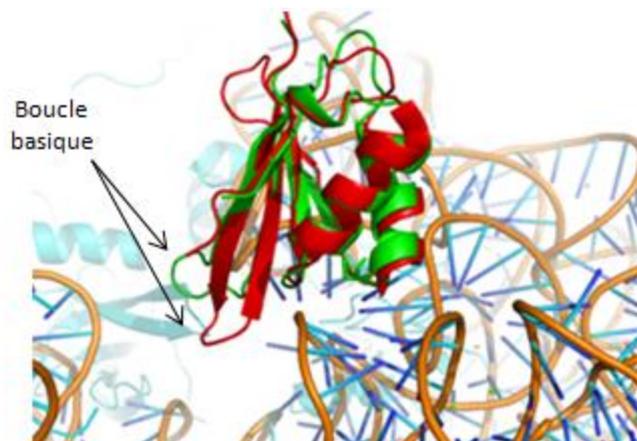


Figure 3.15 : Modélisation de la structure d'eIF1 (rouge) et d'aIF1 de *M. jannaschii* (vert) sur la sous-unité 40S du ribosome

L'éventuelle implication de ces résidus dans la liaison du facteur au ribosome doit dans un premier temps être examinée. Pour cela, afin de vérifier l'implication de la boucle dans sa globalité, tous les acides aminés sont mutés en alanine. La purification de ce mutant, appelé aIF1\*-polyAla est réalisée de la même façon que décrite précédemment (en utilisant toutefois une étape supplémentaire de purification sur colonne à exclusion de taille). Le rendement de marquage étant de 40%, une concentration initiale de 50nM de facteur est suffisante pour obtenir des mesures d'anisotropie avec un rapport fluorescence/bruit élevé.

A la fin du titrage, correspondant à un ajout de 103nM de 30S:ARNm, aucune augmentation de l'anisotropie n'est observée. Le mutant est incapable de se lier à la 30S. Cela indique qu'au minimum un acide aminé se trouvant dans la boucle 30-35 est fortement impliqué dans la liaison au complexe 30S:ARNm

#### 3.1.4.1.2 Caractérisation du mutant aIF1\*-YA<sub>33</sub>

Afin d'étudier plus précisément le rôle de chaque acide aminé contenu dans la boucle 30-35, des mutations ponctuelles de trois résidus de cette boucle ont été réalisées. Un premier mutant de cette boucle, appelé aIF1\*-YA<sub>33</sub> est obtenu après mutation de la tyrosine 33 en alanine. Le rendement de marquage de cette protéine modifiée est de 26%. Il permet d'obtenir une intensité de fluorescence correcte pour une

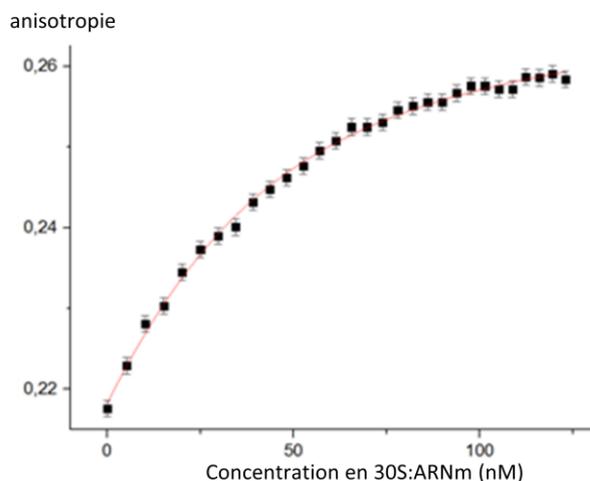


Figure 3.16 : Titrage d'aIF1\*-YA<sub>33</sub> (50nM) avec 30S:ARNm

concentration d'aIF1\*-YA<sub>33</sub> de 50nM.

Le titrage est réalisé par des ajouts successifs de 1.7µL d'une solution à 672nM de 30S et 1340nM d'ARNm. Le résultat est présenté figure 3.16 ci-contre.

La courbe obtenue a une allure hyperbolique. En utilisant les mêmes relations que pour le titrage mettant en jeu aIF1\* et la 30S, nous déterminons le K<sub>d</sub> entre aIF1\*-YA<sub>33</sub> et le complexe 30S:ARNm à  $11.7 \pm 2.8$  nM. **La tyrosine 33 est donc impliquée la formation du complexe aIF1:30S:ARNm et sa substitution diminue l'affinité d'un facteur au moins dix.** Malgré tout, l'affinité mesurée pour le mutant aIF1\*-YA<sub>33</sub> reste mesurable et importante.

#### 3.1.4.1.3 Caractérisation du mutant aIF1\*-KA<sub>35</sub>

Un second mutant de cette boucle 30-35 d'aIF1, correspondant à la mutation de la lysine 35 en alanine (aIF1\*-KA<sub>35</sub>) a été préparé. Le rendement de marquage pour cette protéine est de 18%. Le titrage est réalisé dans les mêmes conditions que pour le titrage de aIF1\*-YA<sub>33</sub>. Le résultat est présenté figure 3.17 ci-contre.

La courbe à l'allure hyperbolique permet de mesurer un K<sub>d</sub> de l'ordre de  $45 \pm 1$  nM. Ce résultat est obtenu après

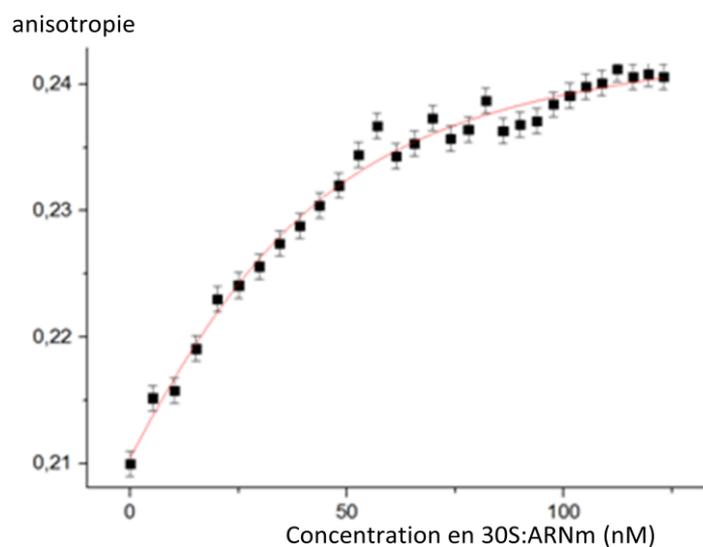


Figure 3.17 : Titrage d'aIF1\*-KA<sub>35</sub> (50nM) avec 30S:ARNm

avoir réalisés 3 fois un titrage identique. L'affinité de ce mutant pour le complexe 30S:ARNm est donc également affectée. **Cet acide aminé contribue à l'affinité par un facteur au moins 10.** Le rendement de marquage de ce variant étant plus faible et le rapport signal/bruit plus faible, l'expérience a été reproduite 4 fois, montrant un profil de courbe identique à chaque expérience.

#### 3.1.4.1.4 Caractérisation du mutant aIF1\*-RA<sub>32</sub>

Le marquage du mutant aIF1\*-RA<sub>32</sub> de la boucle 30-35 correspondant à une mutation de l'arginine 32 en alanine a été réalisé avec un rendement de 24%. Le titrage est réalisé dans les mêmes conditions que pour les deux titrages précédents, et le résultat est présenté figure ci-dessous. Comme pour la forme aIF1\*, la courbe obtenue présente une phase d'accroissement linéaire, puis un plateau de

l'anisotropie lorsque la concentration de 30S:ARNm atteint 70nM (voir figure 3.18 ci-contre). Tout comme pour aIF1\*, cette courbe suggère que le mutant aIF1\*-RA<sub>32</sub> se lie au complexe 30S:ARNm avec une constante de dissociation (Kd) très inférieure à 50nM. Cependant, comme pour aIF1\*, afin d'estimer le Kd d'aIF1\*-RA<sub>32</sub> au complexe 30S:ARNm, la courbe obtenue est traitée avec le logiciel OriginPro. Le traitement de notre courbe conduit à

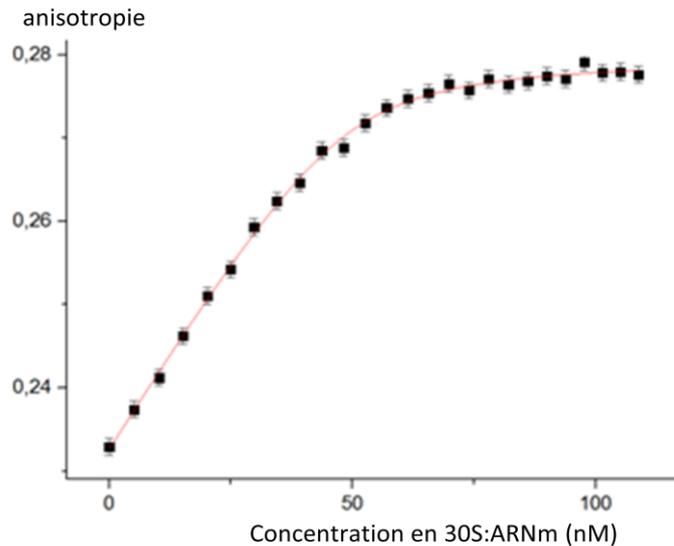


Figure 3.18 : Titrage d'aIF1\*-RA<sub>32</sub> (50nM) avec 30S:ARNm

un Kd entre aIF1\*-RA<sub>32</sub> avec le complexe 30S:ARNm de  $2.4 \pm 0.5$ nM. Cette valeur doit être considérée avec précaution au vu du Kd estimé et de la concentration en enzyme utilisée.

**Finalement, ces résultats indiquent que l'arginine 32 n'est pas impliquée dans la liaison d'aIF1 au complexe 30S:ARNm.**

Une mutation radicale de tous les résidus 30 à 35 composant la boucle basique d'aIF1 entraîne l'incapacité d'aIF1 de se lier à la sous-unité 30S. Les expériences d'anisotropie de fluorescence réalisées avec les mutants ponctuels de cette boucle ont permis de révéler l'implication des résidus Y<sub>33</sub> et K<sub>35</sub> dans la liaison d'aIF1 au complexe 30S:ARNm, le résidu R<sub>32</sub> ne semblant pas impliqué.

Par ailleurs, la détermination des constantes de dissociation des mutants étudiés nous permettra de définir des conditions expérimentales adéquates pour étudier la fonction de ces résidus dans la formation du complexe de démarrage en utilisant la méthode du toeprint.

#### 3.1.4.2 Rôle de l'extrémité N-terminale dans la liaison au complexe 30S:ARNm : caractérisation du mutant aIF1\*-ΔN

Les structures cristallographiques ne permettent pas de révéler la structure de l'extrémité N-terminale des facteurs Tt-eIF1 et Mj-aIF1. Cependant, celle-ci pourrait jouer un rôle dans la fonction d'aIF1 lors du démarrage de la traduction. Nous avons donc voulu tester son implication dans la liaison d'aIF1 au ribosome. Pour cela, les 20 premiers résidus d'aIF1 ont été enlevés par mutagenèse

dirigée du gène, puis la protéine purifiée a été marquée avec du Mdcc. Le marquage a été effectué avec un rendement de 48%. Le titrage est réalisé dans les mêmes conditions que précédemment. Les résultats sont présentés dans la figure 3.19 ci-contre. La courbe de titrage obtenue présente un profil identique à celui des courbes obtenues avec les mutants aIF1\* et aIF1\*-RA<sub>32</sub>, avec une première phase d'accroissement

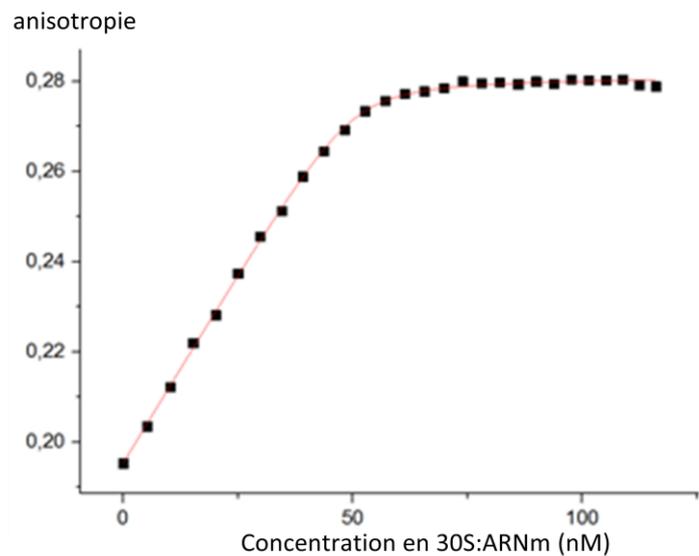


Figure 3.19 : Titrage d'aIF1\*-ΔN (50nM) avec 30S:ARNm

linéaire, et une seconde phase de plateau de l'anisotropie lorsque la concentration de 30S:ARNm atteint 55nM (voir figure ?). Ainsi, le mutant aIF1\*-ΔN possède un Kd très petit devant la concentration en aIF1\*-ΔN (50nM) pour sa liaison avec le ribosome et forme toujours un complexe 1:1 avec la 30S. Comme pour aIF1\* et aIF1\*-RA<sub>32</sub>, la courbe est traitée avec le logiciel OriginPro et conduit à un Kd de  $0.7 \pm 0.1$ nM. Cette valeur doit elle aussi être considérée avec précaution au vu du Kd estimé et de la concentration en enzyme utilisée.

Comme cela pouvait être supposé compte-tenu de la localisation du dernier résidu visible du côté N-terminal dans les structures de Tt-eIF1 et Mj-aIF1 (cf figure 1), la queue N-terminale n'est pas impliquée dans la liaison au complexe 30S:ARNm. Par la suite, nous étudierons le rôle de l'extrémité N-terminale en utilisant la méthode du toeprint.

## 3.2 Etude de la formation et de la stabilité du complexe de démarrage sur l'ARNm

### 3.2.1 Principe du « toeprint »

L'analyse d'un ARN par extension d'amorce est une technique permettant l'étude des séquences des ARNs, de leurs structures ou encore des interactions entre un ARNm et des protéines.

La technique par empreinte du ribosome est une variation de cette technique, alors appelée analyse de l'inhibition de l'extension d'amorce d'un ARN par le complexe de démarrage de la traduction, permettant l'étude de la formation d'un complexe de démarrage de la traduction sur un ARNm et l'implication des différents acteurs dans la liaison du ribosome à ce dernier (Hartz D et al., 1988).

La technique est basée sur le fait que la transcription inverse, catalysée par une transcriptase inverse (ou reverse transcriptase, RT), d'un ARNm peut être mise en pause ou arrêtée lorsque cette dernière rencontre un complexe ribosomique lié spécifiquement à l'ARNm.

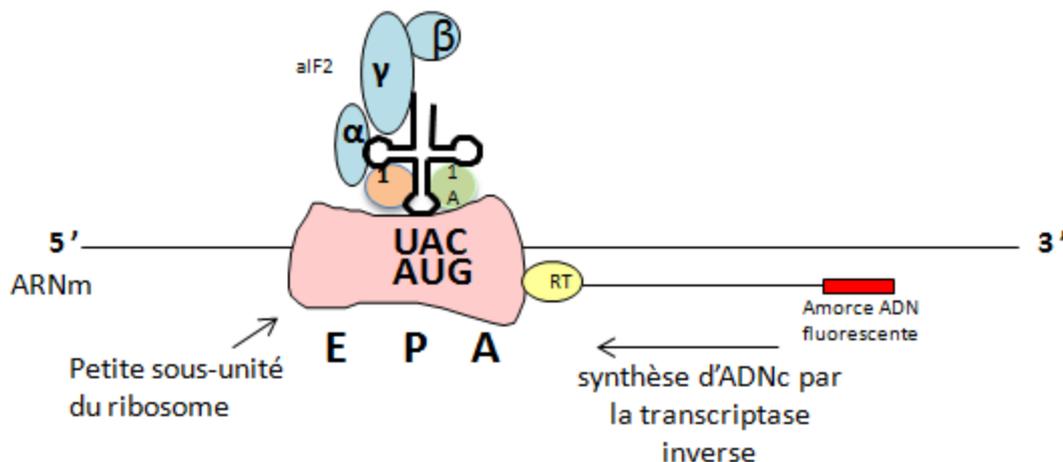


Figure 3.20 : Principe général du toeprinting

Une amorce ADN fluorescente (\*), de séquence complémentaire à la séquence de l'ARNm située en 3' du codon de démarrage de la traduction est hybridée à cet ARNm purifié. Ce complexe hybride ARNm-amorce ADN\* est alors mis en présence de ribosomes, des différents acteurs du démarrage de la traduction, ainsi que d'une RT. L'arrêt prématuré de la RT sur l'ARNm est appelé toeprint ou empreinte du ribosome. L'arrêt de la RT peut facilement être localisé sur l'ARNm en faisant migrer sur un gel de séquence le brin d'ADNc à côté d'une réaction de séquençage de l'ADN utilisé comme matrice de l'ARNm réalisée en utilisant de la même amorce fluorescente.

Ce toeprint permet de repérer la position du ribosome sur l'ARNm modèle selon la combinaison de facteurs de démarrage utilisée, notamment en présence ou en absence d'aIF1, et d'évaluer la stabilité de ces différents complexes. Le but est de connaître la contribution d'aIF1 et de ses différents mutants dans la stabilisation du complexe de démarrage sur l'ARNm.

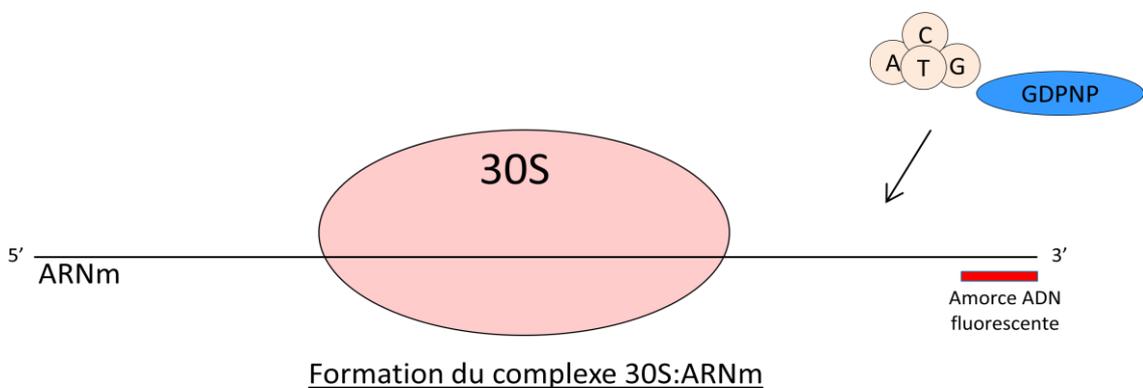
Chaque expérience de toeprint se déroulera comme suit :

- Tout d'abord, l'hybridation de l'amorce ADN sur l'ARNm modèle est réalisée dans un tampon d'hybridation à pH 7.5 contenant du  $\beta$ -mercaptoéthanol et 600mM de chlorure d'ammonium. La solution contenant 6.6 picomoles d'amorce ADN fluorescente et 1.1 picomoles d'ARNm modèle est chauffée 3 minutes à 60°C afin de favoriser l'hybridation, puis de l'acétate de magnésium est ajouté pour atteindre la concentration de 60mM adéquate pour le ribosome.

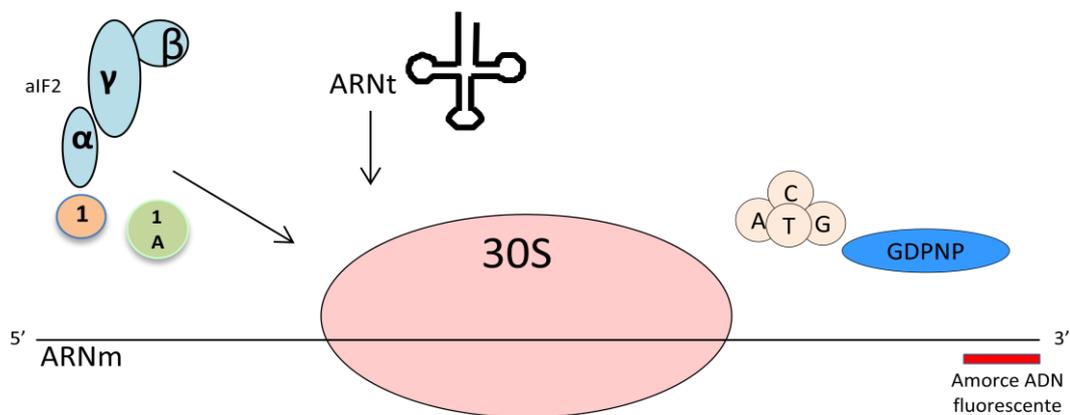


- L'étape suivante consiste en la préparation d'un prémix contenant la sous-unité du ribosome (53nM), des dNTP (416 $\mu$ M) (nécessaires à l'élongation de l'amorce par la reverse transcriptase), du GDPNP (555 $\mu$ M) (assurant la saturation du facteur aIF2 pendant l'expérience) et le mélange d'hybridation contenant l'ARNm à 70nM. A cette étape, la petite sous-unité du ribosome est saturée en ARNm compte-tenu du Kd estimé par anisotropie de fluorescence à 3.6nM.

Le prémix est alors chauffé 5 min à 51°C afin de favoriser la liaison entre l'ARNm et la 30S.

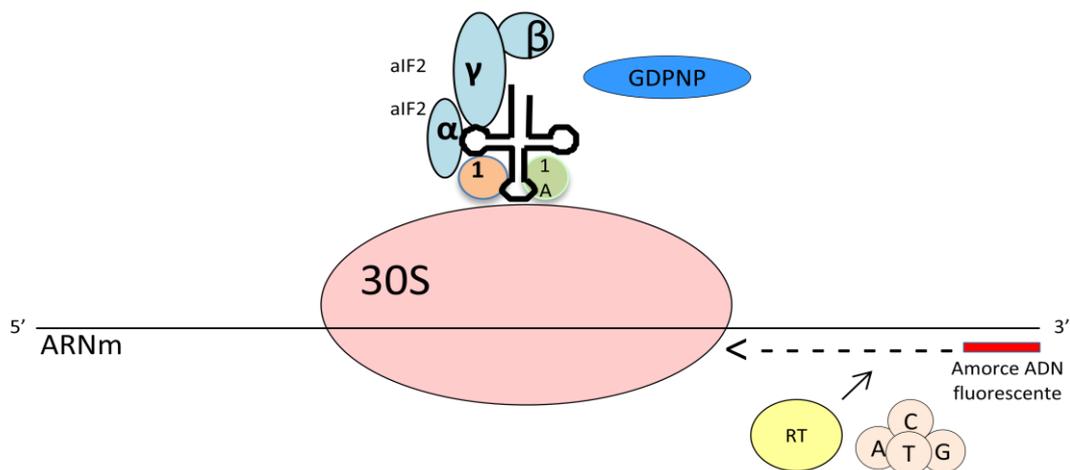


- Finalement les facteurs de démarrage sont mélangés à ce pré-mix. Le mélange final de toeprint contient alors des dNTP, du GDPNP, de l'acétate de magnésium, du tampon d'hybridation, la petite sous-unité du ribosome (25.10nM), l'ARNm modèle (33.15nM), les facteurs de démarrage souhaités (263.1nM chacun) et l'ARNt d'intérêt (1052nM). Dans ce mélange, et compte-tenu des différents Kd déterminés par anisotropie de fluorescence et connus (le Kd du complexe eIF2:ARNt est de 90nM pour l'ARNt<sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>, Yatime L et al., 2004), chaque sous-unité 30S est vraisemblablement saturée en ARNm, en facteur de démarrage et en ARNt.



Formation du complexe de démarrage de la traduction d'intérêt

- Une première incubation de ce mélange 10 minutes à 51°C est réalisée. 0.25U de reverse transcriptase (RT) est alors ajoutée puis les tests sont incubés 15 minutes à 51°C. Nous avons utilisés la RT de l'Avian Myeloblastosis Virus (AMV) plutôt que celle du Moloney Murine Leukemia Virus (MMLV). Cette dernière est en effet stoppée par des complexes ribosomiques liés de manières moins stables à l'ARNm (P. Fechter et al., 2009).



Rétrotranscription de l'ARNm par la reverse transcriptase

	Concentration dans le mélange réactionnel final de toeprint
30S	25,1nM
ARNm	33,15nM
facteur	263,1nM
ARNt	1052nM

Après l'arrêt de la réaction, les échantillons peuvent être déposés sur le gel de séquence et l'image du gel de séquence obtenue (Licor : Base ImageIR Data Collection version 4.0) est traitée à l'aide du logiciel Image J.

Dans cette partie, nous étudierons l'effet des facteurs aIF1 et dans une moindre mesure aIF1A dans la liaison du ribosome à l'ARNm, ainsi que leur participation éventuelle dans la sélection du bon codon de démarrage et de l'ARNt initiateur.

Toutes les expériences de toeprint sont réalisées avec un fragment d'ARNm de 600 bases codant pour le gène aEF1A de *P. abyssi* synthétisé par transcription *in vitro* (voir Matériels et Méthodes). Cet ARNm contient une séquence Shine-Dalgarno forte avec 9 bases complémentaire de l'extrémité 3' de l'ARNr 16S. L'ARNm modèle synthétique de 26 bases utilisé lors des expériences d'anisotropie de fluorescence correspond à la région 5' de ce messenger :

Séquence de l'ARNm aEF1A WT au voisinage du codon de démarrage:  
 5' CAGAAATTTGGAGGTGATTTAAATGCCAAAGGAGA ..... 3'  
                                   SD                                  Met Pro  
   1   2

### 3.2.2 Conditions de stabilisation de la sous-unité 30S du ribosome sur l'ARNm.

#### 3.2.2.1 Stabilisation du ribosome sur l'ARNm

La réaction de toeprint permet de visualiser sur un gel de séquence la position d'arrêt de la petite sous-unité du ribosome sur l'ARNm lors du démarrage de la traduction en présence ou en absence des différents acteurs du démarrage. Le protocole de toeprint est adaptable selon le nombre et la composition du jeu de facteurs de démarrage testés.

Afin de pouvoir observer une éventuelle variation des signaux d'arrêts (toeprinting) de la 30S sur l'ARNm, il est dans un premier temps nécessaire de connaître le signal de toeprinting obtenu sur l'ARNm en présence de chacun des facteurs en complexe avec la 30S. Lorsque le toeprint est réalisé avec l'ARNm et la sous-unité 30S seule, aucun signal d'arrêt n'est observé. La 30S seule n'est pas suffisamment stabilisée sur l'ARNm. Lorsque les facteurs aIF1A ou aIF2 sont ajoutés au milieu réactionnel, de très faibles bandes sont visibles sur le gel de polyacrylamide-urée au niveau des signaux d'arrêt de toeprint, et à l'inverse une bande forte de signal de transcription totale est visible, indiquant un signal d'arrêt et une stabilité très faible des complexes 30S:aIF1A et 30S:aIF2 sur l'ARNm aux concentrations testées. Au contraire, lorsque le facteur aIF1 seul est complexé à la sous-unité 30S, trois bandes peuvent être observées sur le gel de séquence, ainsi qu'un signal de transcription totale plus faible que précédemment (voir figure 3.21). Ces trois bandes correspondent chacune à un arrêt de la RT reflétant la stabilisation du ribosome sur l'ARNm (voir figure 3.22).

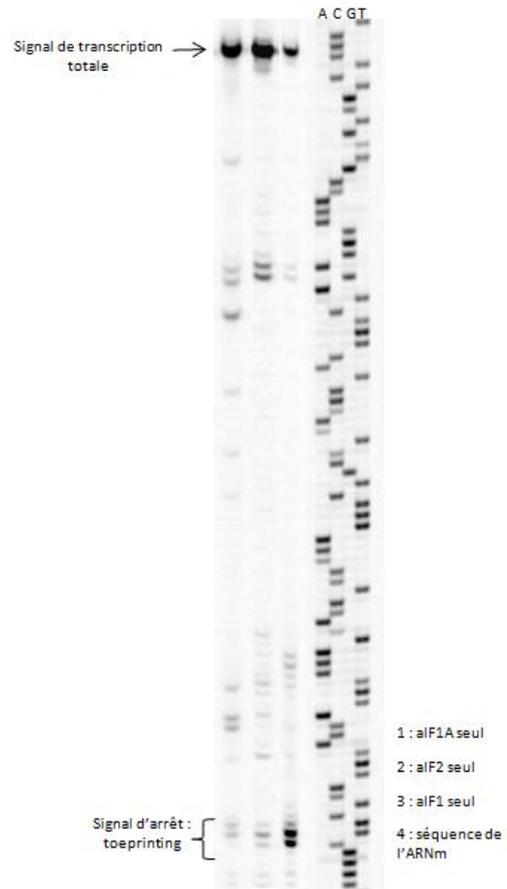


Figure 3.21 : Migration des réactions de toeprint sur gel polyacrylamide-urée : signaux de toeprinting obtenu avec les complexes 30S:aIF1, 30S:aIF1A ou 30S:aIF2 sur l'ARNm modèle

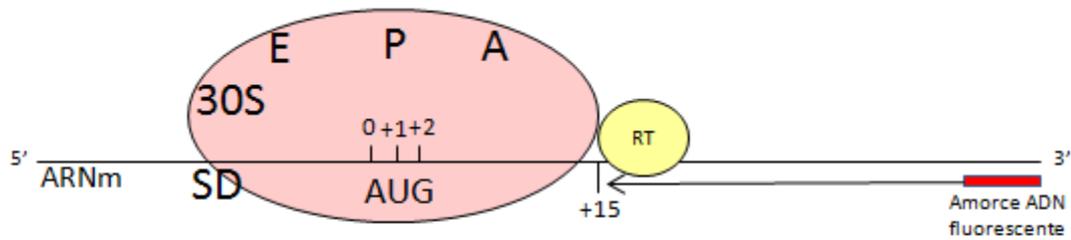


Figure 3.22 : Modèle de la méthode du toeprint. La reverse transcriptase termine la synthèse de l'ADNc quand elle rencontre la 30S stable sur l'ARNm.

Il est alors possible de connaître la position du ribosome sur l'ARNm en alignant les différents arrêts observés et la séquence de l'ADN correspondant. La position 0 est définie pour la base A du codon de démarrage de la traduction sur l'ARNm (voir figure 3.22)

Lorsque aIF1 est complexé au complexe 30S:ARNm, les signaux de toeprint sont observés en positions +12, et plus fortement en position +13 et +14 (voir figure 3.23).

La force du signal de toeprint rapportée au signal de rétro-transcription jusqu'à l'extrémité 5' de l'ARNm (signal de transcription totale) est une mesure quantitative de la formation et de la stabilité du complexe de démarrage. L'image du gel, obtenue sur un séquenceur automatique à détection de fluorescence, est traitée à l'aide du logiciel ImageJ qui permet ici de déterminer l'amplitude des signaux de toeprint et de rétro-transcription totale. Le logiciel calcule ensuite les aires sous les pics correspondant aux signaux, puis le rapport  $\text{aire}_{\text{toeprint}}/\text{aire}_{\text{totale}}$  peut être calculé ( $\text{aire}_{\text{totale}} = \text{aire}_{\text{toeprint}} + \text{aire}_{\text{transcription totale}}$ ) (voir figure 3.24).

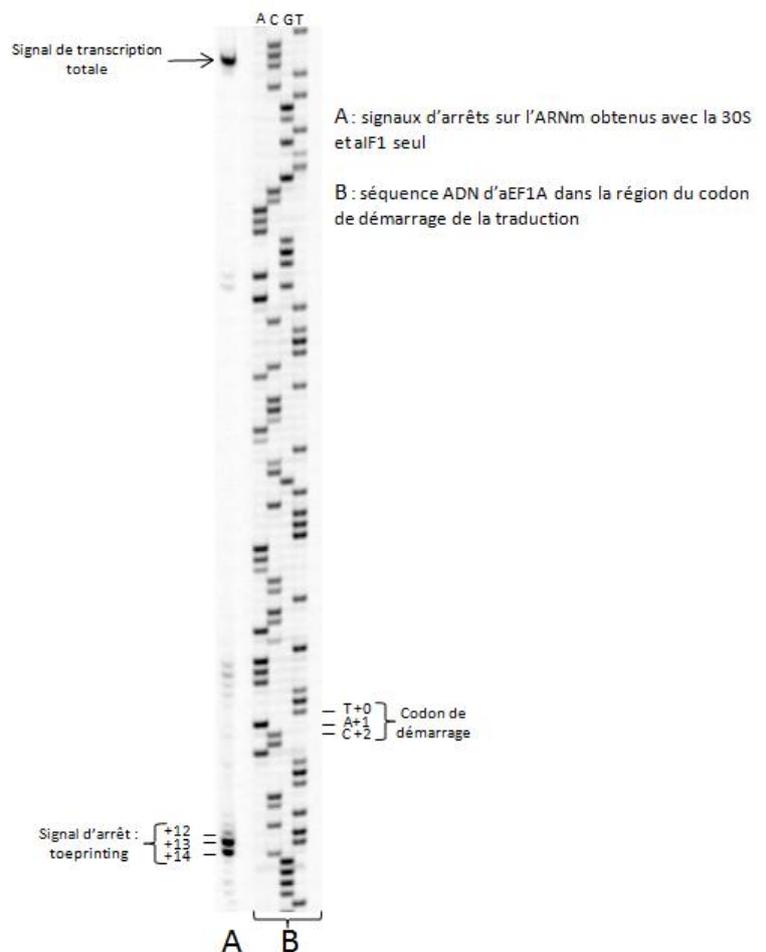
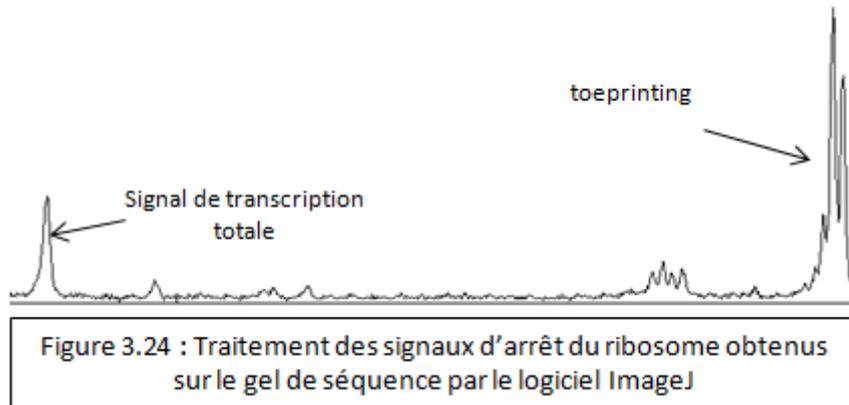
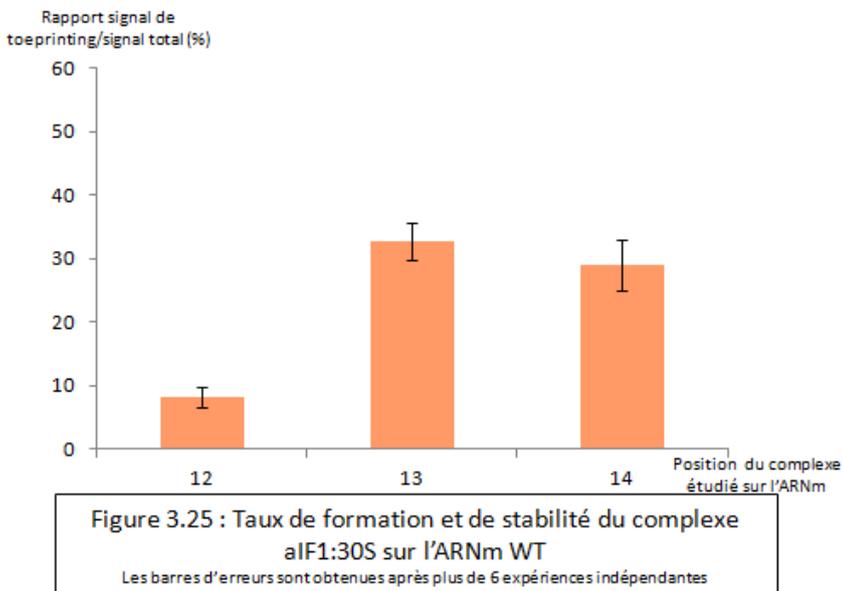


Figure 3.23 : Migration d'une expérience de toeprint et de la séquence ADN d'aEF1A sur gel de polyacrylamide-urée : toeprinting obtenus avec le complexe 30S:aIF1



Le taux de stabilité du ribosome sur l'ARNm lorsque seul le facteur aIF1 est présent dans le milieu peut donc être visualisé comme suit dans la figure 3.25 :



Les barres d'erreurs correspondent à l'écart-type calculé à partir de 6 expériences réalisées dans les mêmes conditions. Lorsque les barres d'erreurs sont présentes, cela signifie que l'expérience a été répétée dans les mêmes conditions au moins 2 fois. Toutes les expériences comprenant le facteur aIF1 WT et l'ARNm WT ont été réalisées plus de 6 fois.

### 3.2.2.2 Caractérisation de la saturation du ribosome en facteur aIF1

Les tests de toeprint sont réalisés avec 10 fois plus de facteur que de 30S (263nM d'aIF1 pour 25.1nM de 30S), ce qui, au vu du Kd estimé par anisotropie de fluorescence (de l'ordre du nanomolaire) est très suffisant pour s'assurer de la saturation de la 30S en facteur aIF1. Cependant, afin de s'assurer que la concentration en aIF1 utilisée pendant le test de toeprint n'influence pas les résultats et que la concentration définie au départ est saturante pour la concentration en ribosome utilisée dans le test, une gamme de concentration en aIF1 est réalisée. Cette gamme varie de 0nM en aIF1 dans le milieu réactionnel à 2105nM, c'est-à-dire 80 fois plus d'aIF1 que de 30S. Le profil de

toeprint observé ainsi que l'intensité des signaux sont invariants à partir de 131.5nM d'aIF1 (voir figure 3.26). Ces résultats peuvent paraître surprenant compte tenu du Kd d'aIF1 pour le complexe 30S:ARNm déterminé par anisotropie de fluorescence (de l'ordre du nanomolaire). Cependant, à cause de régions riches en résidus basiques, des liaisons aspécifiques entre aIF1 et l'ARNm peuvent se former,

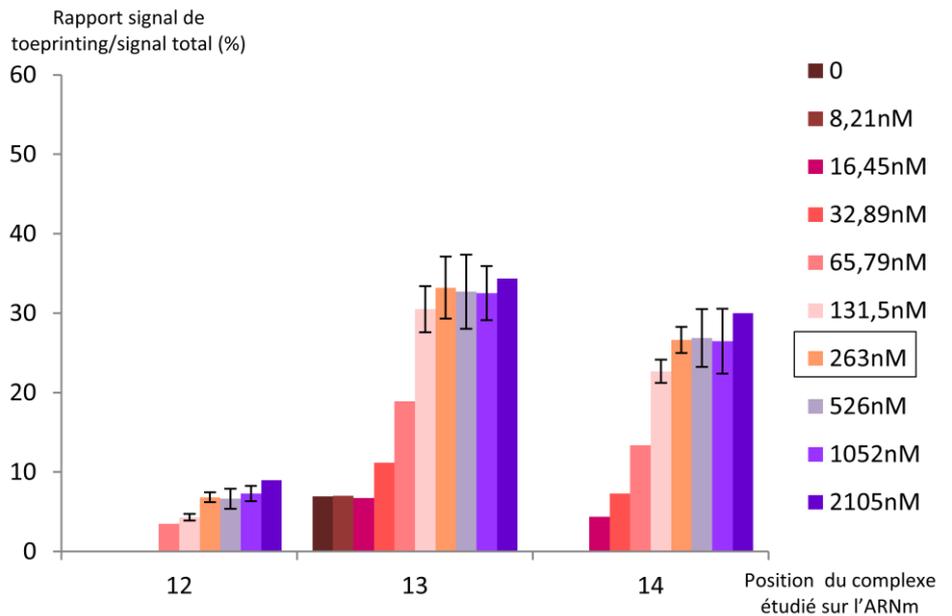


Figure 3.26 : Toeprint avec le complexe 30S:aIF1 et l'ARNm WT. Gamme de concentration d'aIF1. Les barres d'erreurs sont obtenues après plus de 2 expériences indépendantes

diminuant ainsi la concentration en facteur disponible pour se lier à la 30S. En effet, contrairement à l'ARNm utilisé en anisotropie (26 bases), l'ARNm de 600 bases utilisé en toeprint n'est pas entièrement couvert par la 30S. aIF1 est donc susceptible de se lier de façon aspécifique à ce dernier. Ainsi, pour assurer la saturation en ribosome, nous utiliserons une concentration en aIF1 de 263nM dans les tests de toeprint.

### 3.2.2.3 Influence de la séquence de l'ARNm dans la formation du complexe aIF1:30S sur l'ARNm

#### 3.2.2.3.1 Influence de la position du codon de démarrage sur l'ARNm

Afin de tester si la position de toeprint du complexe 30S:aIF1 sur l'ARNm observée est due à la position du codon de démarrage de la traduction (voir figure ci-dessous), une version mutée de l'ARNm modèle est réalisée. Ce nouvel ARNm contient en première position un codon codant pour une valine (Val1) et le codon de démarrage est positionné en position 2 (Met2). Il est appelé ARNm Val1Met2 dans la suite des expériences.

Séquence de l'ARNm aEF1A WT centrée sur la séquence SD :

```

5' CAGAAATTTGGAGGTGATTTAAATGCCAAAGGAGA ..... 3'
      SD           Met Pro
      1           2

```

Séquence de l'ARNm aEF1A Val1Met2 centrée sur la séquence SD :

```

5' CAGAAATTTGGAGGTGATTTAAAGTAATGAAGGAGA ..... 3'
      SD           Val Met
                    1  2

```

Lorsque l'expérience de toeprint est réalisée dans les mêmes conditions que précédemment avec le facteur aIF1 seul et l'ARNm Val1Met2, le signal de toeprinting observé est similaire à celui observé lorsque l'ARNm WT était utilisé (voir figure 3.27).

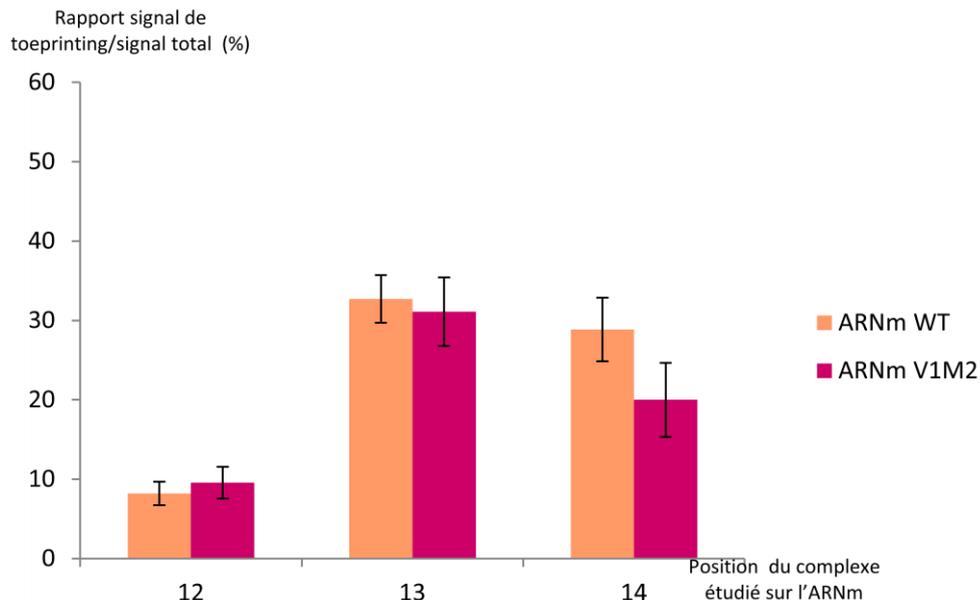


Figure 3.27 : Comparaison du taux de formation et de stabilité du complexe aIF1:30S sur l'ARNm WT ou Val1Met2

Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes

Ces résultats nous indiquent que l'empreinte observée avec la sous-unité 30S et le facteur aIF1 ne dépend pas de la position du codon méthionine par rapport à la séquence SD. La présence d'une séquence SD forte sur l'ARNm pourrait jouer un rôle dans la formation de ce complexe 30S:aIF1 sur l'ARNm.

Afin de conforter ces résultats, des expériences similaires ont été réalisées avec des ARNm mutés dans leur codon 1 et/ou 2 :

Séquence de l'ARNm aEF1A WT au voisinage du codon de démarrage:

```

5' CAGAAATTTGGAGGTGATTTAAATGCCAAAGGAGA ..... 3'
      SD           Met Pro
                    1  2

```

Séquence de l'ARNm aEF1A Val1Met2 centrée sur la séquence SD :

```

5' CAGAAATTTGGAGGTGATTTAAAGTAATGAAGGAGA ..... 3'
      SD           Val Met
                    1  2

```

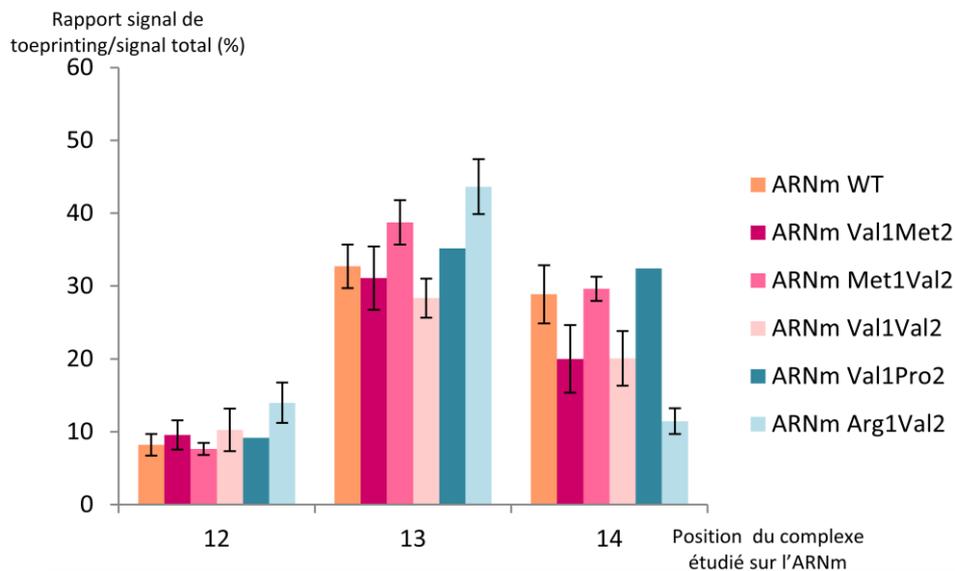
Séquence de l'ARNm aEF1A Met1Val2 centrée sur la séquence SD :  
 5' CAGAAATTTGGAGGTGATTTAAATGGTAAAGGAGA ..... 3'  
                           SD                          Met Val  
   1 2

Séquence de l'ARNm aEF1A Val1Val2 centrée sur la séquence SD :  
 5' CAGAAATTTGGAGGTGATTTAAGTGTAAAGGAGA ..... 3'  
                           SD                          Val Val  
   1 2

Séquence de l'ARNm aEF1A Val1Pro2 centrée sur la séquence SD :  
 5' CAGAAATTTGGAGGTGATTTAAGTACCAAGGAGA ..... 3'  
                           SD                          Val Pro  
   1 2

Séquence de l'ARNm aEF1A Arg1Val2 centrée sur la séquence SD :  
 5' CAGAAATTTGGAGGTGATTTAAAGGTAAAGGAGA ..... 3'  
                           SD                          Arg Val  
   1 2

Les résultats sont présentés dans la figure 3.28 ci-dessous.



**Figure 3.28 : Toeprint avec le complexe 30S:aIF1 et l'ARNm WT ou des ARNm mutés dans leur codon 1 et/ou 2.**

Les barres d'erreurs avec les ARNm WT et Val1Met2 sont obtenues après 6 expériences indépendantes

Les barres d'erreurs avec les autres ARNm sont obtenues après plus de 2 expériences indépendantes

Les résultats obtenus avec l'ARNm Met1Val2 montrent un même taux de formation et de stabilité du complexe 30S:aIF1 sur les ARNm variants qu'avec l'ARNm WT. La nature du codon situé en aval du

codon 1 méthionine n'influe pas sur la formation ni sur la stabilité du complexe 30S:aIF1 sur l'ARNm. Les résultats obtenus avec les ARNm Val1Val2, Val1Pro2 et Arg1Val2 sont eux aussi similaires à ceux obtenus avec l'ARNm WT. On peut ainsi en déduire que la présence d'un codon méthionine n'est pas nécessaire à la formation d'un complexe 30S:aIF1 sur un ARNm possédant une séquence SD forte, et que la nature des codons situés après cette séquence SD n'a de même pas d'influence sur la stabilité de ce complexe. Ainsi, la seule présence d'une séquence SD pourrait être suffisante pour permettre la stabilisation du complexe 30S:aIF1 sur l'ARNm.

Pour confirmer le fait que la stabilisation de l'ARNm sur la 30S par aIF1 est indépendante du codon méthionine, nous avons mesuré par anisotropie de fluorescence l'affinité d'aIF1\* pour un complexe 30S:ARNm obtenu en utilisant un ARNm dans lequel un codon AUG est remplacé par un codon CAU. La courbe obtenue est équivalente à la courbe obtenue avec l'ARNm modèle possédant un codon méthionine. Le traitement de cette courbe avec le logiciel OriginPro conduit à un Kd de  $0.8 \pm 0.9$  nM. Ainsi, nous confirmons que la liaison d'aIF1 au complexe 30S:ARNm ne nécessite pas de codon méthionine.

### 3.2.2.3.2 Influence de la séquence SD

Afin d'étudier l'implication de la séquence SD, différents ARNm sont utilisés. Le premier est un ARNm correspondant à l'ORF d'aIF2 $\gamma$  contenant une séquence SD différant seulement de 2 bases (notées -8 et -5) par rapport à l'ARNm modèle WT tout en conservant la séquence GGAGG. La séquence SD est alors complémentaire de l'ARNr 16S par 7 bases et non plus par 9. Le second ARNm est un variant de l'ARNm WT portant seulement la séquence consensus GGAGG, appelé par la suite ARNm MutSD. Les bases mutées sont numérotées -5,-6,-7,-8 (la base A du codon ATG étant la base 0). Les bases qui diffèrent de la séquence SD de l'ARNm d'aEF1A sont en gras ci-dessous.

Séquence de l'ARNm aIF2 $\gamma$  centrée sur la séquence SD :

```

5' AACACTTCTGGAGGCGAGAGCCATGGGAGAGAAG ..... 3'
           SD  -8  -5      Met Gly
  
```

Séquence de l'ARNm aEF1A MutSD centrée sur la séquence SD :

```

5' CAGAAATTAGGAGGACTATTAATGCCAAGGAGA .... 3'
           SD  -8-7-6-5      Met Pro
  
```

Lorsque l'ARNm aIF2 $\gamma$  est utilisé, un signal de toeprinting du complexe aIF1:30S équivalent au signal obtenu avec l'ARNm WT est constaté (voir figure 3.29 ci-dessous).





bactérie, appelé ARNt<sub>f</sub><sup>Met</sup>. Le second est l'ARNt initiateur de *P. abyssi* produit par surexpression dans une souche d'*E. coli*, appelé ARNt<sub>i</sub><sup>Met</sup>. Il faut noter que le niveau d'expression de cet ARNt chez *E. coli* est faible. L'ARNt initiateur de *P. abyssi* est donc seulement partiellement purifié (aminoacybilité mesurée à 900 picomoles/unité d'absorbance à 260nm). De plus, les modifications post-transcriptionnelles de cet ARNt ne sont pas connues. Il est donc aussi possible qu'il soit produit sous forme hypomodifiée chez *E. coli*. Finalement, le troisième ARNt initiateur est l'ARNt<sub>f</sub><sup>Met</sup> muté dans ses bases 1 et 72, appelé ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>. Cet ARNt est un bon mime de l'ARNt<sub>i</sub> d'archée (voir chapitre 1). Nous avons réussi à le produire en grande quantité et à le purifier très efficacement grâce à la souche d'*E. coli* modifiée (PalΔMetZWV:KanRΔMetY, pBSTnav2A<sub>1</sub>-U<sub>72</sub>) décrite dans le chapitre 1.

Pour chaque expérience de toeprint, les ARNt initiateurs sont préalablement méthionylés (voir Matériels et Méthodes).

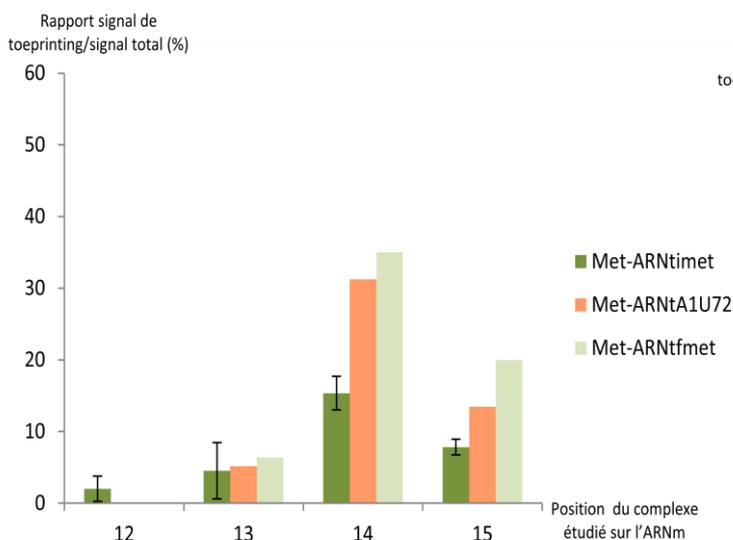


Figure 3.31 : Taux de formation et de stabilité du complexe ARNt:30S sur l'ARNm WT

Les barres d'erreurs sont obtenues après 2 expériences indépendantes

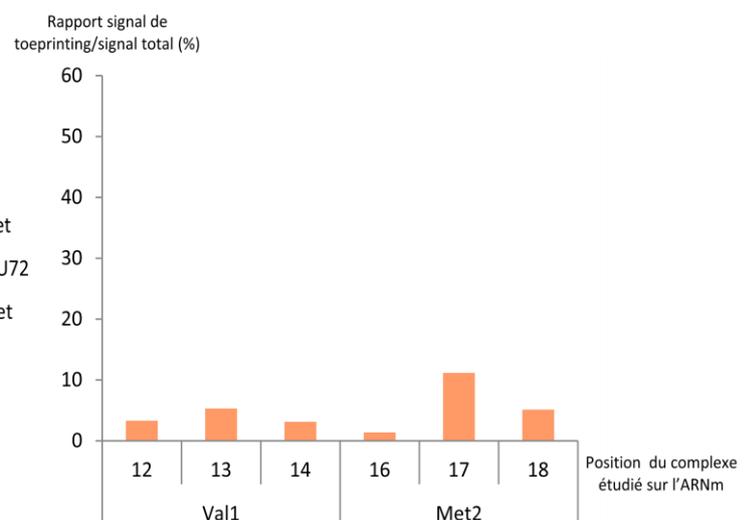


Figure 3.32 : Taux de formation et de stabilité du complexe ARNtA<sub>1</sub>U<sub>72</sub>:30S sur l'ARNm Val1Met2

Le profil des signaux de toeprint observé avec ces ARNts est équivalent pour l'ARNt<sub>f</sub><sup>Met</sup> et pour l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> (figure 3.31 ci-dessus). Le signal obtenu avec l'ARNt<sub>i</sub><sup>Met</sup> de *P. abyssi* est toutefois plus faible. Cet effet est vraisemblablement lié à une moins grande pureté de cet ARNt et à sa possible contamination par des ARNt élongateurs. Les signaux de toeprinting se trouvent principalement en position +14 et +15, et un signal très faible est observé en position +13 et +12. Au vu de ces résultats et compte tenu de la pureté des préparations, les expériences suivantes seront réalisées avec l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>.

Finalement, si le codon AUG est déplacé en position 2, comme avec l'ARNm Val1Met2, le signal majoritaire est déplacé en position +17 et +18 (voir figure 3.32 ci-dessus). Les signaux obtenus sont

plus faibles qu'avec l'ARNm WT. Cet effet est vraisemblablement dû au fait que le codon méthionine en position 2, plus éloigné de la séquence SD, est moins favorable à la stabilité du complexe 30S:ARNt sur l'ARNm que lorsqu'il est en position 1.

Ainsi, l'empreinte en position +14 et +15 est attribuée au signal correspondant à un complexe où l'ARNt initiateur est apparié au codon AUG dans le site P. De même, lorsqu'un ARNm possédant un codon AUG en position 2 est utilisé, les positions +17 et +18 sont attribuées au signal de l'ARNt<sub>i</sub> apparié au site P. (voir figure ci-dessus). Ceci démontre la capacité de l'ARNt initiateur à provoquer le positionnement de la sous-unité 30S avec le codon de démarrage au site P.

### 3.2.3 Etude de différents complexes de démarrage sur la stabilité du complexe 30S:ARNm

Tableau récapitulatif des termes employés pour le toeprint dans la suite des expériences :

Abbréviations	Acteurs présents dans le complexe
CT (complexe ternaire)	Met-ARNt <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub> :aIF2
CT + aIF1	Met-ARNt <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub> :aIF2:aIF1
CI (complexe d'initiation)	Met-ARNt <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub> :aIF2:aIF1:aIF1A
CI sans aIF1	Met-ARNt <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub> :aIF2:aIF1A
CI sans aIF1A	Met-ARNt <sub>f</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub> :aIF2:aIF1

#### 3.2.3.1 Le complexe CT:30S:ARNm

En présence du facteur aIF2 et du Met-ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>, le signal de toeprint est augmenté et est restreint spécifiquement aux positions +14 et +15 (en violet dans la figure 3.33 ci-dessous).

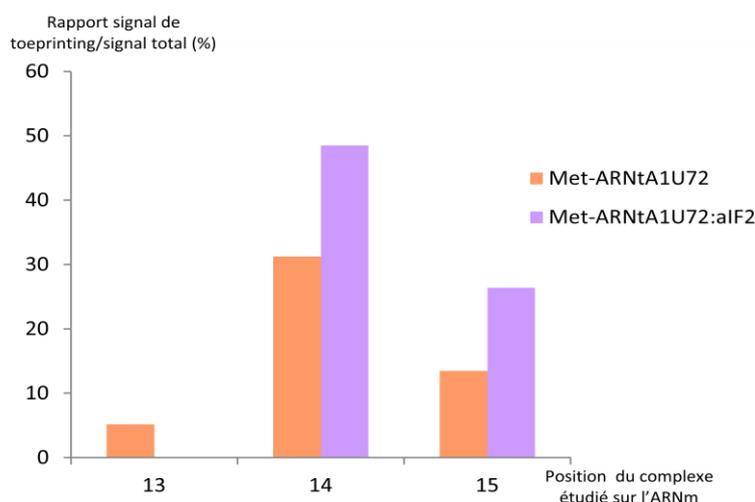


Figure 3.33 : Effet d'aIF2 sur la stabilité du complexe Met-ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>:30S sur l'ARNm WT

L'effet d'aIF2 dépend de l'état de méthionylation de l'ARNt initiateur. Si l'ARNt<sub>i</sub> n'est pas méthionylé, seulement un très faible effet d'aIF2 est observé sur la stabilisation. Ce résultat est en accord avec le fait que l'affinité d'aIF2 pour l'ARNt dépend très fortement de son aminoacylation avec de la méthionine (Yatime L. et al., 2004). En effet, l'affinité d'un ARNt initiateur « valylé » à aIF2 est diminuée de plus de 2 ordres de grandeur par rapport à celle mesurée pour un ARNt initiateur méthionylé (Kd > 50000nM pour Val-ARNt<sub>f</sub><sup>Met</sup> contre Kd = 90nM pour le Met-ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>)

De façon similaire, lorsque l'ARNm Val1Met2 est utilisé, le signal de toeprint est très augmenté et quasi restreint aux positions +17 et +18 lorsque la 30S est mise en présence du complexe Met-ARNtA<sub>1</sub>U<sub>72</sub>:aIF2 (voir figure 3.34 ci-dessous).

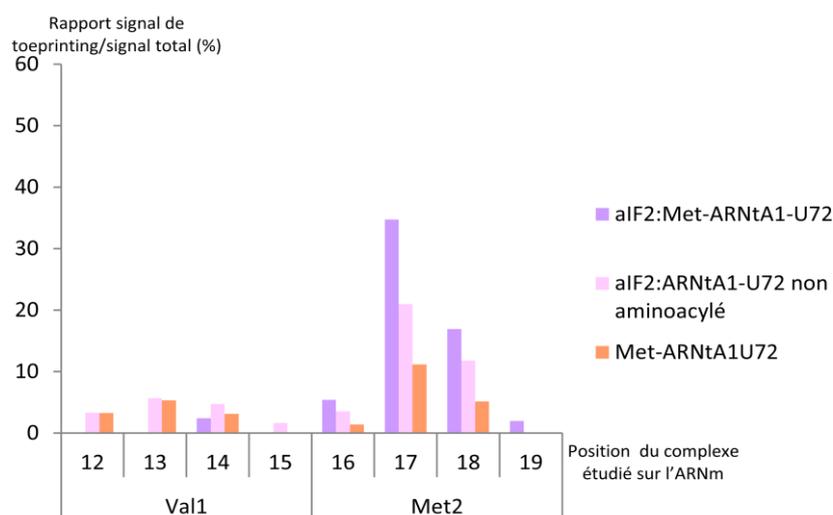


Figure 3.34 : Effet d'aIF2 sur la stabilité du complexe Met-ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>:30S sur l'ARNm WT

### 3.2.3.2 Variation de la concentration en ARNt

Les expériences de toeprint impliquant l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> ont été réalisées avec des concentrations dans le milieu expérimental de 1052nM d'ARNt<sub>i</sub> pour 25.1nM de 30S et pour 263nM d'aIF2, ce qui, au vu des Kd connus de l'ARNt<sub>i</sub> pour la 30S et pour aIF2 devrait nous permettre de nous placer en condition saturante. Cependant, et afin de vérifier que la concentration en ARNt<sub>i</sub> utilisée dans le test n'influe pas sur les résultats, la réaction de toeprint précédente impliquant la 30S, l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> et le facteur aIF2 a été reproduite avec une gamme de concentration de l'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> allant de 52.26 à 2105 nM d'ARNt<sub>i</sub> dans le milieu réactionnel. Le profil et l'intensité de toeprinting obtenu ne varie pas en fonction de la concentration. En se plaçant à 1052nM d'ARNt<sub>i</sub> dans le milieu réactionnel, nous assurons donc pleinement la saturation du ribosome en CT.

### 3.2.3.3 Le complexe Met-ARNtA<sub>1</sub>U<sub>72</sub>:aIF2:aIF1:30S:ARNm (CI sans aIF1A)

Lorsque CT:30S:ARNm est mis en présence du facteur aIF1, le signal de toeprint en +15 est fortement diminué et des arrêts apparaissent en position +12 et plus fortement en +13 (en violet dans la figure 3.35 ci-contre).

Ainsi, la présence d'aIF1 sur le CT:30S:ARNm suffit à le déstabiliser. On peut proposer qu'aIF1 provoque un changement de conformation de la sous-unité 30S, passant alors d'une

conformation stable centrée sur le codon AUG à une conformation plus dynamique dans laquelle l'ARNt initiateur ne serait plus complètement apparié de façon permanente au codon de démarrage. Cette expérience sera répétée afin de diminuer les barres d'erreurs.

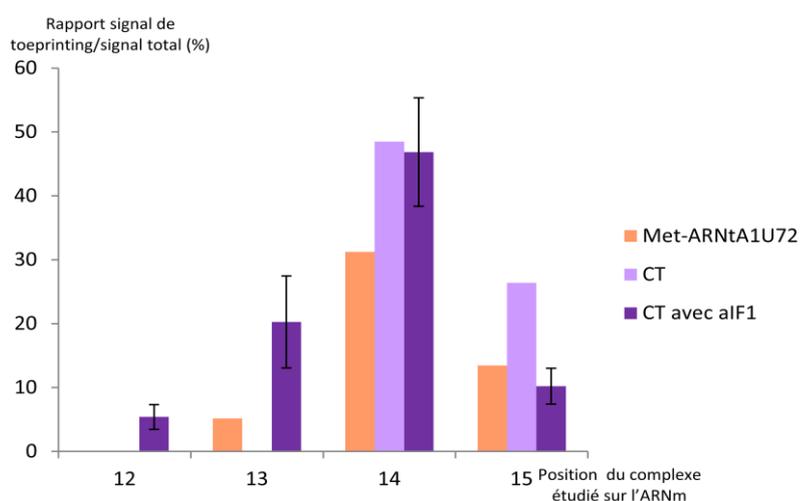


Figure 3.35 : Taux de formation et de stabilité des différents complexes de démarrage sur l'ARNm WT  
Les barres d'erreurs sont obtenues après 2 expériences indépendantes

3.2.3.4 III.2.4 : Signal de toeprinting du complexe de démarrage complet : Met-ARNtA<sub>1</sub>U<sub>72</sub>:aIF2:aIF1:aIF1A:30S:ARNm (CI)

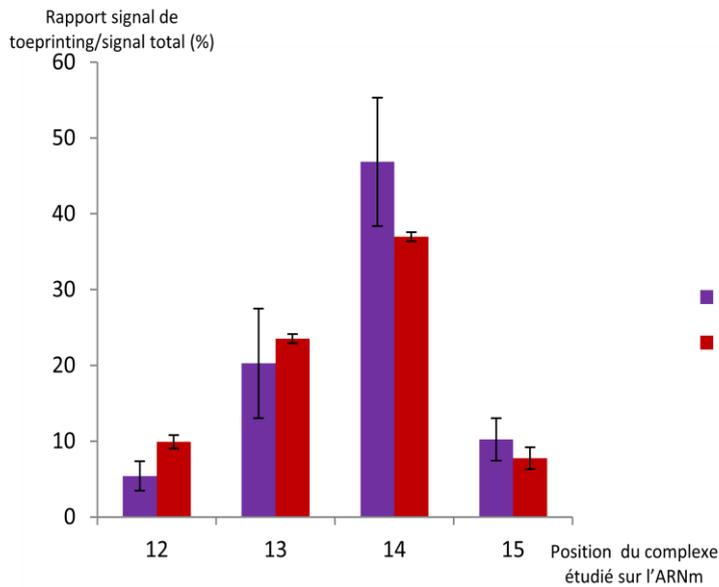


Figure 3.36 : Influence de la présence de tous les facteurs de démarrage dans la formation du complexe de démarrage sur l'ARNm WT

Les barres d'erreurs pour CI sans aIF1A sont obtenues après 2 expériences indépendantes  
Les barres d'erreurs pour CI sont obtenues après 6 expériences indépendantes

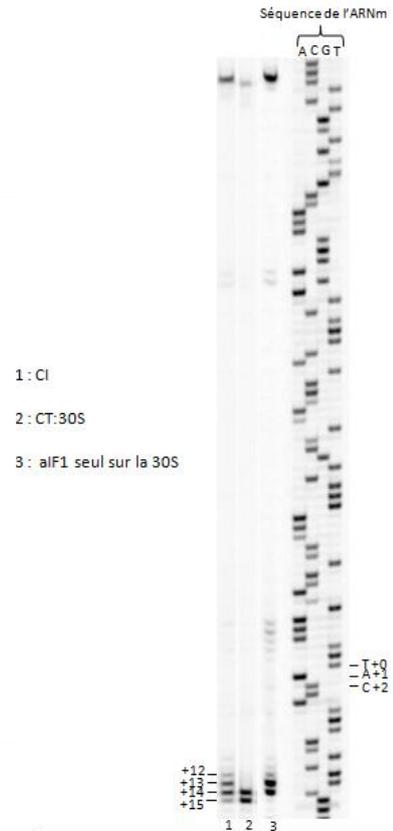


Figure 3.37 : Migration de réactions de toeprinting sur gel de polyacrylamide-urée : signaux de toeprinting obtenus avec le CI (1), le CT (2) ou aIF1 seul (3)

Lorsque le facteur aIF1A est ajouté au milieu réactionnel, permettant ainsi la formation du complexe total du démarrage de la traduction, l'effet conformationnel provoqué par aIF1 est encore augmenté. On constate en effet une diminution plus importante du signal de toeprinting en position +15 et +14, accompagnée de l'augmentation du signal en position +13 et +12 (voir figures 3.36 et 3.37 ci-dessus). Ainsi, en présence d'aIF1, le facteur aIF1A renforce le rôle déstabilisateur d'aIF1 du complexe CT:30S:ARNm centré sur le codon AUG.

Finalement, lorsque le facteur aIF1 est retiré du milieu, on observe un retour du profil de toeprinting en positions +14 et +15, validant ainsi l'effet observé (voir figure 3.38 ci-dessous).

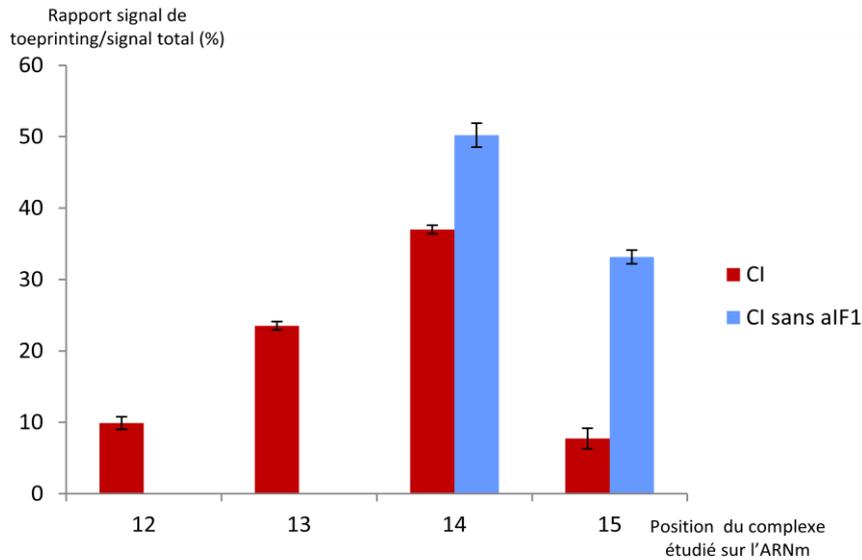


Figure 3.38 : Influence d'aIF1 dans la formation et la stabilité du complexe de démarrage sur l'ARNm WT

Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes

Les mêmes expériences permettant de mettre en évidence l'influence d'aIF1 sur la conformation du complexe de démarrage ont été réalisées avec l'ARNm Val1Met2.

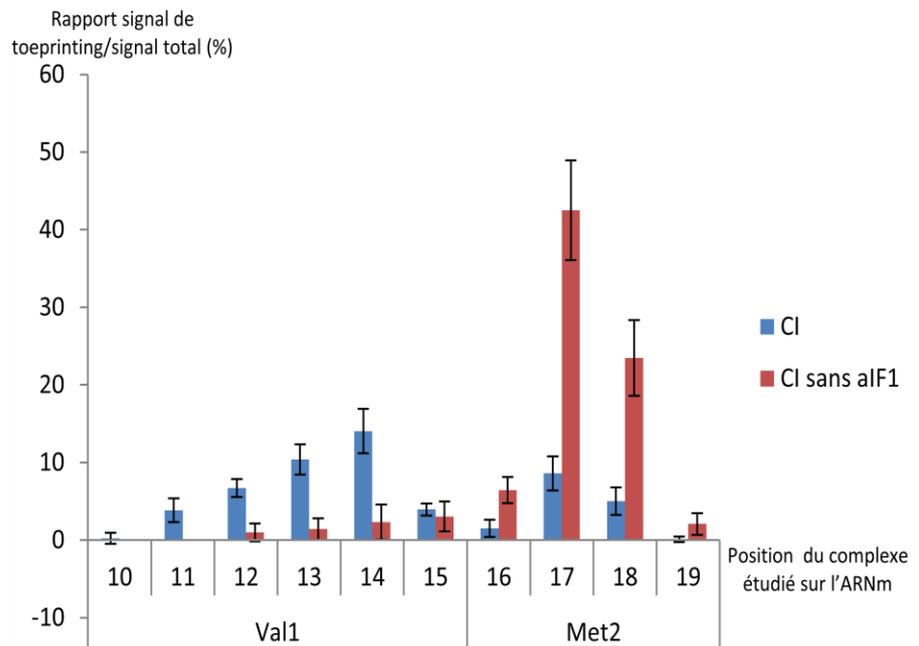
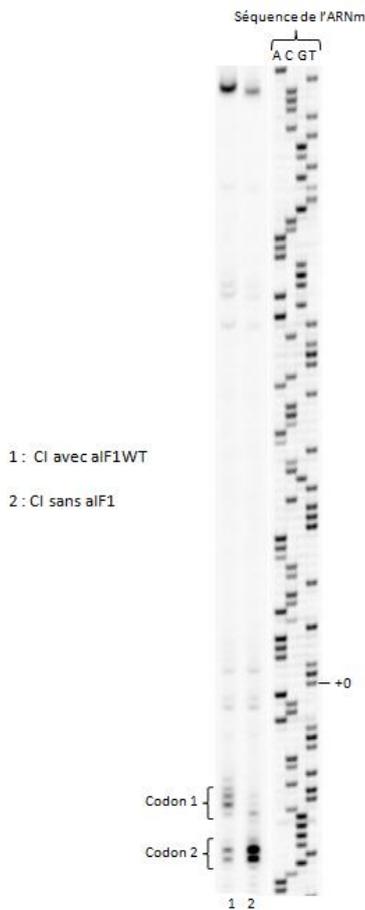


Figure 3.40 : taux de formation et de stabilité du CI sur l'ARNm Val1Met2

Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes

Figure 3.39: migration de réactions de toeprint sur gel de polyacrylamide-urée : signaux de toeprinting obtenu avec le CI avec et sans aIF1 sur l'ARNm Val1Met2

Les profils des toeprinting pour chaque expérience sont similaires à ceux réalisés avec l'ARNm WT (voir figures 3.39 et 3.40 ci-dessus). Lorsqu'aIF1 est absent (CI sans aIF1) du milieu expérimental, les signaux de toeprinting sont centrés sur les positions +17 et +18 indiquant un ajustement du complexe ribosomal sur le codon Met2 dans une conformation permettant l'appariement de l'ARNt initiateur au codon AUG centré dans le site P. L'addition d'aIF1A renforce le signal aux positions +17 et +18. Enfin de la même manière que précédemment, en présence d'aIF1, le signal en position +17 et +18 diminue fortement et le profil de toeprinting est déstabilisé vers les positions +13 et +14, indiquant un changement de conformation de la sous-unité 30S sur l'ARNm. Ces résultats confirment l'action déstabilisante d'aIF1 sur le complexe centré sur le codon de démarrage, et ce quelle que soit la position (1 ou 2) de ce codon par rapport à la séquence SD.

**Pour conclure, ces expériences montrent qu'aIF1 stabilise l'ARNm sur la 30S. Cet effet est indépendant de la position du codon, mais il dépend de la force de la séquence Shine-Dalgarno. En présence de l'ensemble des facteurs, le complexe de démarrage a un caractère plus dynamique avec l'anticodon de l'ARNt qui n'est pas apparié de façon stable au codon de démarrage de l'ARNm. Cet effet disparaît en absence d'aIF1, qui est donc responsable de cette dynamique accrue. Une telle conformation dynamique est propice à une activité d'édition des mauvais complexes de démarrage (impliquant un mauvais codon) par aIF1.**

#### **3.2.4 Etude du rôle des résidus de la boucle basique d'aIF1**

Nous avons pu mettre en évidence par des expériences de titrage en anisotropie de fluorescence l'implication de deux résidus de la boucle de résidus basiques d'aIF1 dans sa liaison à la petite sous-unité du ribosome. La mutation de ces résidus en alanine diminue l'affinité d'aIF1 à la sous-unité mais elle n'annule pas complètement cette liaison.

Afin de pouvoir étudier au mieux le rôle de chaque résidu ou domaine pouvant jouer un rôle dans le démarrage de la traduction, chaque variant étudié en anisotropie de fluorescence est replacé dans le contexte de la séquence « sauvage » d'aIF1. En effet, ces derniers contiennent chacun une double mutation CS63, KC46 dans leur séquence nécessaire à la liaison du fluorophore. La liaison du fluorophore n'étant pas nécessaire pour les expériences de toeprint, les variants d'intérêts d'aIF1 sont obtenus par mutagenèse dirigée par PCR du gène codant pour aIF1 sauvage. Ainsi, seul le résidu d'intérêt est muté.

Cependant, nous avons malgré tout vérifié que la double mutation CS63KC46 d'aIF1 nécessaire aux expériences d'anisotropie de fluorescence ne modifiait pas la fonction d'aIF1 observée par la

méthode de toeprint. Les résultats obtenus avec aIF1\* sont tout à fait similaires à ceux obtenus avec aIF1WT. Nous confirmons ainsi que la variant d'aIF1 utilisé en anisotropie de fluorescence n'est pas modifié dans ses fonctions.

### 3.2.4.1 Etude du rôle de l'arginine 32

Les études par anisotropie de fluorescence n'ont pas révélé de rôle majeur de l'arginine 32 d'aIF1 dans la liaison de ce dernier au complexe 30S:ARNm ( $K_d \sim 2.4\text{nM}$ ). Cependant, le positionnement du résidu homologue dans la structure 3D de Tt-eIF1 lié à la 40S suggère que ce résidu puisse quand même jouer un rôle dans la formation du complexe de démarrage de la traduction sur l'ARNm

#### 3.2.4.1.1 Le complexe aIF1-RA32:30S sur l'ARNm WT

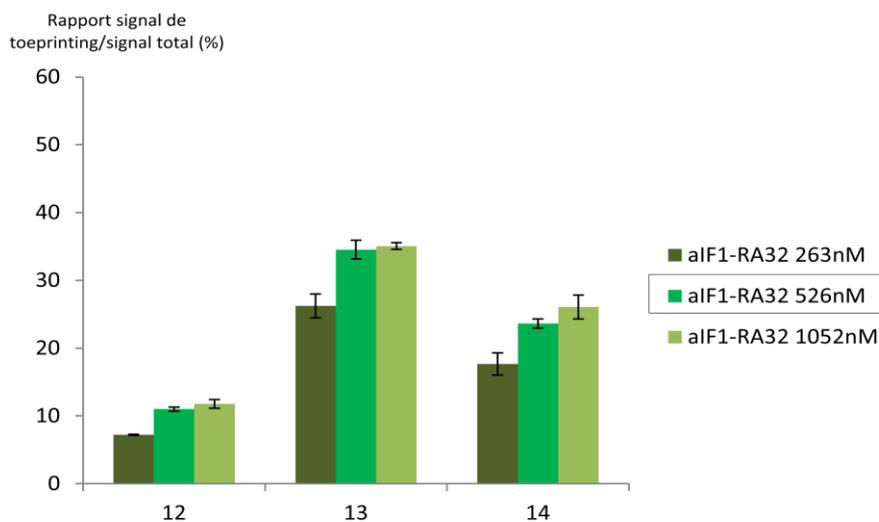


Figure 3.41 : Toeprint avec le complexe 30S:aIF1-RA<sub>32</sub> et l'ARNm WT. Gamme de concentration en aIF1-RA<sub>32</sub>  
*Les barres d'erreurs avec aIF1-RA<sub>32</sub> sont obtenues après 2 expériences indépendantes*

Les expériences sont réalisées dans les mêmes conditions qu'avec le facteur aIF1 WT, c'est-à-dire avec un mélange réactionnel contenant la sous-unité 30S du ribosome, l'ARNm et le facteur aIF1 dans lequel l'arginine 32 a été préalablement mutée en alanine (aIF1-RA32).

Afin de s'assurer d'être dans des conditions assurant la saturation du ribosome en facteur aIF1-RA32, une gamme de concentration a été réalisée. Les résultats sont visibles dans la figure 3.41 ci-dessus.

Nous observons une légère variation du signal d'arrêt en fonction de la concentration en aIF1-RA32 dans le milieu.

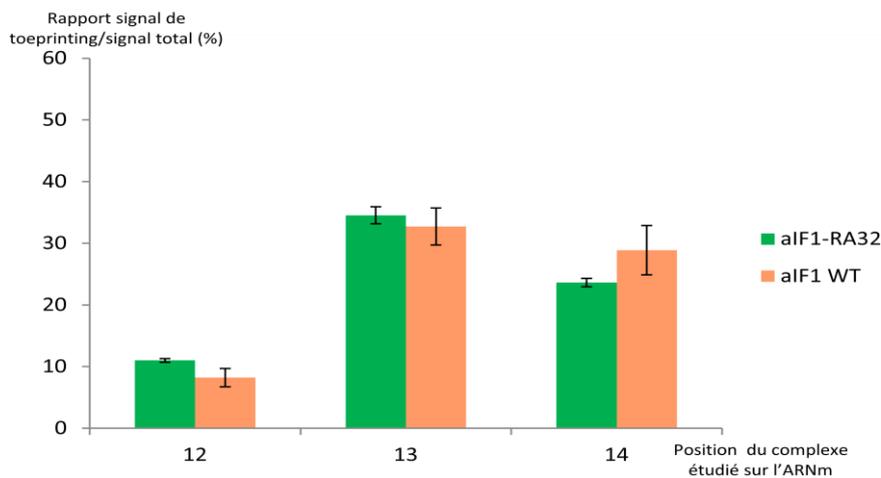


Figure 3.42: Toeprint avec le complexe 30S:aIF1-RA<sub>32</sub>, 30S:aIF1WT sur l'ARNm WT.

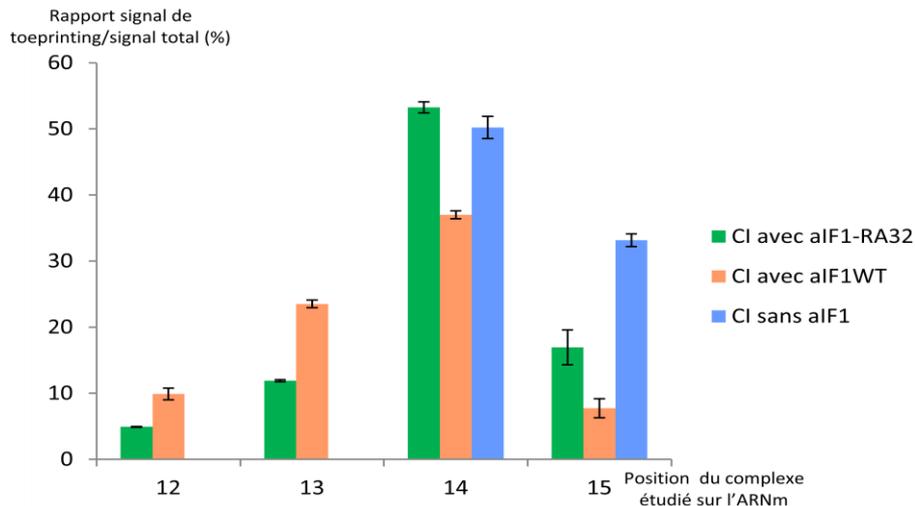
*Les barres d'erreurs avec aIF1-RA<sub>32</sub> sont obtenues après 2 expériences indépendantes  
Les barres d'erreurs avec aIF1 WT sont obtenues après plus de 6 expériences indépendantes*

Pour la suite des expériences avec ce mutant, et afin de se placer dans les conditions optimales de saturation de la 30S, nous nous placerons à 526nM d'aIF1-RA32 dans le milieu réactionnel. Comme pour aIF1WT, on peut proposer que l'utilisation d'un ARNm de 600 bases conduit à des interactions non spécifiques d'aIF1 et à la baisse de sa concentration disponible pour la liaison à la 30S.

Les profils d'empreintes obtenus avec les complexes aIF1-RA32:30S et aIF1WT:30S sont identiques et centrés sur les positions +13 et +14 (voir figure 3.42 ci-dessus). Il apparaît donc qu'en présence ou en absence d'arginine 32, les complexes 30S:ARNm:aIF1 adoptent une conformation similaire à condition que la saturation en facteur soit effectivement assurée.

#### 3.2.4.1.2 Le CI avec aIF1-RA32

Afin de pouvoir observer l'effet de l'arginine 32 dans la formation du complexe de démarrage de la traduction, les expériences de toeprint sont menées avec le CI et aIF1-RA32 dans les conditions de saturation en facteur préalablement définies.

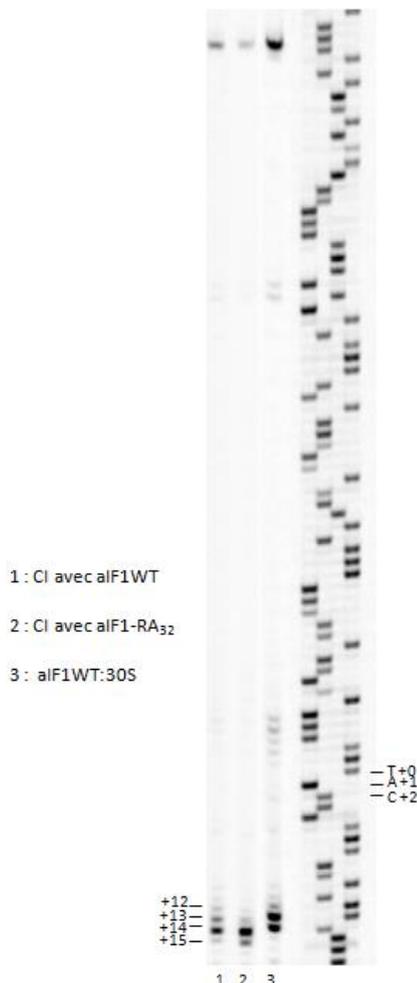


**Figure 3.43: Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1 ou d'aIF1-RA<sub>32</sub>**

*Les barres d'erreurs avec aIF1-RA<sub>32</sub> sont obtenus après 2 expériences indépendantes  
Les barres d'erreurs avec aIF1 WT et sans aIF1 sont obtenus avec plus de 6 expériences indépendantes*

Cette fois, le profil de toeprinting observé avec le CI et aIF1-RA32 est différent de celui observé avec aIF1WT. En effet, le profil obtenu avec aIF1-RA32 n'est plus centré sur les positions +13 et +14, mais est déplacé vers les positions +14 et +15 (en vert dans la figure 3.43 ci-dessus).

En présence de tous les facteurs, le variant aIF1-RA32 perd en partie sa capacité à déstabiliser le complexe de démarrage centré sur le codon AUG (voir la figure 3.44 ci-contre). Le profil de toeprinting avec le CT et aIF1-RA32 se rapproche alors du profil obtenu lorsqu'aIF1 n'est plus dans le milieu (CT sans aIF1) (en bleu dans la figure 3.43 ci-dessus), avec cependant une intensité d'empreinte moins élevée en +14 et +15. Une plus forte proportion d'ARNt initiateur serait donc appariée au codon de démarrage au niveau du site P du ribosome.



**Figure 3.44 : Migration de réactions de toeprint sur gel de polyacrylamide-urée : signaux de toeprinting obtenu avec le CI aIF1WT (1) ou aIF1-RA<sub>32</sub> (2) ou avec aIF1 seul (3)**

**Le résidu arginine 32 de la boucle basique d'aIF1 est donc impliqué dans le rôle déstabilisateur d'aIF1 du complexe de démarrage centré sur le codon AUG.**

### 3.2.4.2 Etude du rôle de la tyrosine 33

Nos études par anisotropie de fluorescence ont révélé l'implication de la tyrosine 33 d'aIF1 dans la liaison de ce dernier à la petite sous-unité du ribosome ( $K_d = 13.2 \pm 0.5\text{nM}$ ). Cependant, comme pour chaque résidu de la boucle basique, la tyrosine 33, de par son emplacement sur la 30S pourrait jouer un rôle plus important dans la fonction d'aIF1 lors du démarrage de la traduction. Une étude plus approfondie de son rôle est donc réalisée par toeprint en utilisant un variant d'aIF1 possédant une alanine à la place de la tyrosine 33 (aIF1-YA33).

#### 3.2.4.2.1 Lle complexe aIF1-YA33:30S sur l'ARNm WT : rôle dans la stabilisation de la 30S sur l'ARNm

Tout comme pour le variant aIF1-RA32, une gamme de concentration allant de 263nM à 2105nM dans le milieu réactionnel d'aIF1-YA33 en complexe avec la 30S a été réalisée. Les signaux d'arrêts obtenus ne varient pas en fonction de la concentration en aIF1-YA33 dans le milieu. D'autre part, les signaux de toeprinting observables avec aIF1-YA33 sont très faibles comparés aux signaux obtenus avec aIF1WT (voir figures 3.45 et 3.46 ci-dessous).

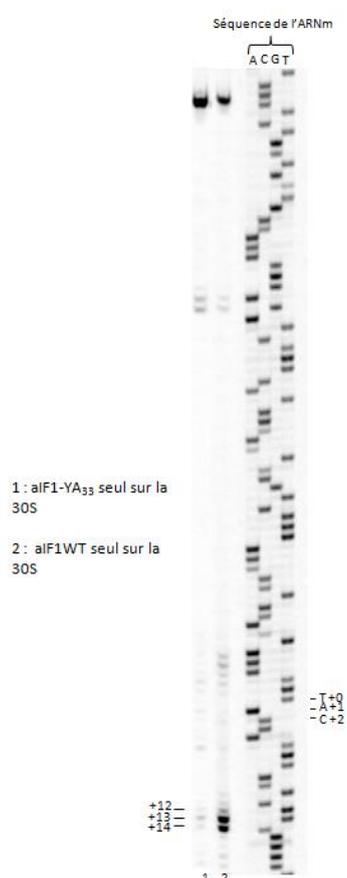


Figure 3.45 : Migration de réactions de toeprint sur gel de polyacrylamide-urée : signaux de toeprinting obtenu avec aIF1WT et aIF1-YA<sub>33</sub> sur la 30S.

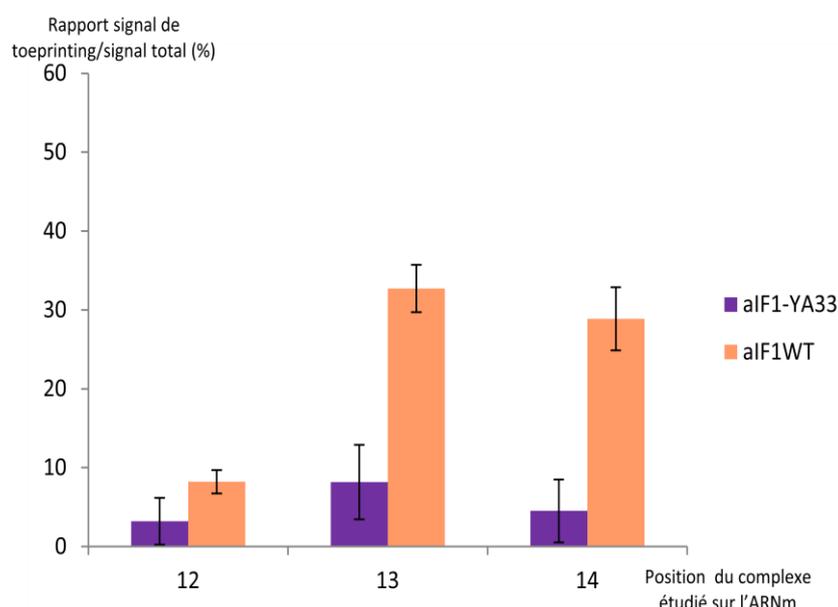


Figure 3.46 : Toeprint avec le complexe 30S:aIF1-YA<sub>33</sub>, 30S:aIF1WT sur l'ARNm WT.

Les barres d'erreurs avec aIF1-YA<sub>33</sub> sont obtenues après 4 expériences indépendantes  
Les barres d'erreurs avec aIF1 WT sont obtenues avec plus de 6 expériences indépendantes

On observe aussi une plus grande intensité du signal de rétrotranscription total de l'ARNm, nous indiquant que la RT est effectivement moins bien arrêtée sur l'ARNm par le ribosome. Même dans des conditions saturantes, le variant aIF1-YA33 a donc perdu sa capacité à stabiliser la sous-unité 30S du ribosome sur l'ARNm et la tyrosine 33 est impliquée de manière très forte dans cette fonction.

De la même manière, lorsque l'ARNm Val1Met2 est utilisé, la stabilisation du ribosome sur l'ARNm reste très faible mais tout de même centrée sur les positions +13 et +14. Ce défaut de stabilisation de la 30S sur l'ARNm reste donc indépendant des positions 1 ou 2 du codon méthionine sur l'ARNm. De façon cohérente, des expériences de titrage d'aIF1\*-YA33 avec la sous-unité 30S seule par anisotropie de fluorescence nous ont permis de constater que le Kd de ce variant pour la 30S était très semblable lorsque le titrage est réalisé avec ou sans ARNm (de 13.2 à 14.3nM respectivement). Ces résultats d'anisotropie révèlent l'absence de coopérativité entre la liaison d'aIF1 et la liaison de l'ARNm lorsque la tyrosine 33 est mutée en alanine.

Ces résultats sont en accord avec ceux obtenus en toeprint. En effet, dans les deux types d'expériences, on peut constater que lorsque la tyrosine 33 est mutée, la stabilisation de la 30S sur l'ARNm est très fortement diminuée.

**Ainsi, nous concluons que la tyrosine 33 est impliquée dans la stabilisation de l'ARNm sur la 30S.**

#### 3.2.4.2.2 Le CI avec aIF1-YA33

Compte tenu des résultats précédents, nous avons choisi de nous placer à la concentration de 263nM en aIF1-YA33 dans la suite des expériences. Le complexe de démarrage a été formé en utilisant aIF1-YA33. Les résultats sont visibles en violet dans la figure 3.47 ci-dessous.

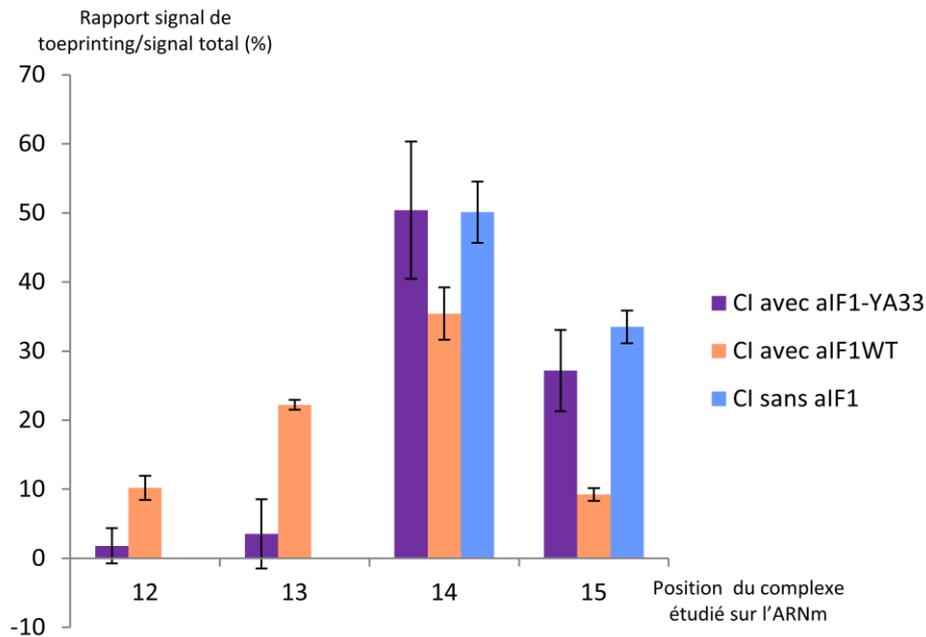


Figure 3.47 : Comparaison des profils de toeprinting obtenus pour le CI en présence d'aIF1WT ou d'aIF1-YA<sub>33</sub> sur l'ARNm WT  
 Les barres d'erreurs avec aIF1-YA<sub>33</sub> sont obtenus après 2 expériences indépendantes  
 Les barres d'erreurs avec aIF1 WT et sans aIF1 sont obtenus avec plus de 6 expériences indépendantes

Tout comme pour le variant aIF1-RA32, le profil de toeprinting obtenu avec le CI et aIF1-YA33 est décalé vers les positions +14 et +15, de manière très similaire au profil obtenu avec le CI sans aIF1. Lorsque la tyrosine 33 est changée en alanine, le facteur aIF1 perd presque en totalité la capacité de déstabiliser le complexe centré sur le codon de démarrage. Ce résidu Y33 est donc très fortement impliqué dans la formation du complexe dynamique précédent l'appariement définitif de l'ARNt au codon AUG.

De façon similaire, lorsque l'ARNm Val1Met2 est utilisé, le profil de toeprinting est décalé préférentiellement vers les positions +17 et +18 comme lorsque le facteur aIF1 est omis de la réaction (CI sans aIF1)(voir figures 3.48 et 3.49 ci-dessous). Ces positions sont définies comme étant les positions sur lesquelles l'ARNt est apparié au codon de démarrage au niveau du site P du ribosome. **Ces résultats nous indiquent cette fois encore que le variant aIF1-YA33 perd en partie sa capacité à déstabiliser le complexe de démarrage, même lorsque celui-ci est déjà dans un contexte moins favorable de stabilisation compte tenu de la position du codon méthionine.**

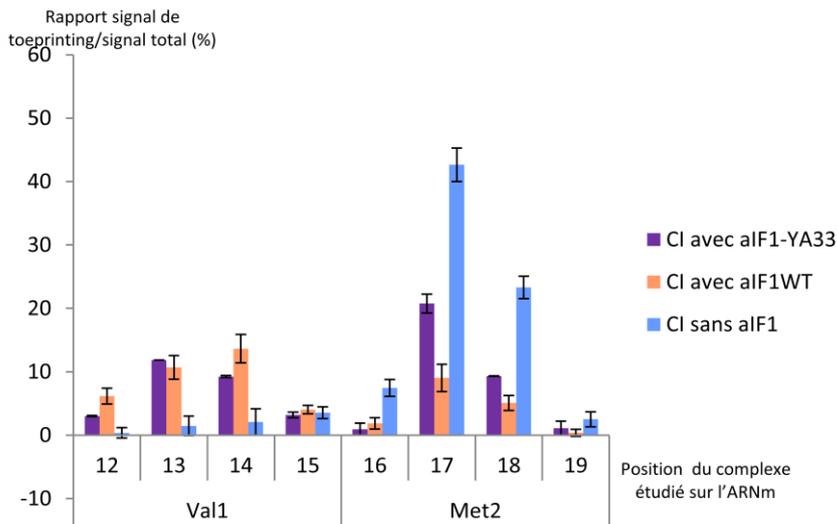


Figure 3.48: Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1WT ou d'aIF1-YA<sub>33</sub> sur l'ARNm Val1Met2

Les barres d'erreurs avec aIF1-YA<sub>33</sub> sont obtenus après 2 expériences indépendantes  
 Les barres d'erreurs avec aIF1 WT et sans aIF1 sont obtenus avec plus de 6 expériences indépendantes

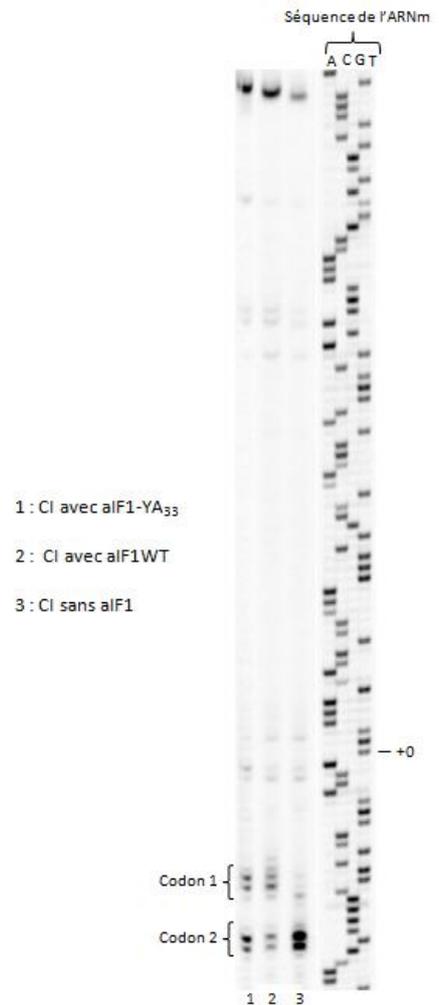


Figure 3.49 : Migration de réactions de toeprint sur gel de polyacrylamide-urée : signaux de toeprinting obtenu avec le CI et aIF1-YA<sub>33</sub> (1) ou aIF1WT (2) et le CI sans aIF1 (3) sur l'ARNm Val1Met2

### 3.2.4.3 Etude du rôle de la lysine 35

L'étude du variant d'aIF1 muté ponctuellement au niveau de la lysine 35 par anisotropie de fluorescence a permis de mettre en évidence la très forte participation de ce résidu dans la liaison d'aIF1 à la 30S (Kd = 45nM pour la liaison avec 30S:ARNm). Pour continuer l'étude, nous avons construit un variant d'aIF1 dans lequel la lysine 35 est remplacé en alanine (aIF1-KA35). Une gamme de concentration de 263 à 2105nM de facteur a été réalisée (voir figure 3.40 ci-dessous). On peut remarquer que l'intensité des signaux de toeprinting obtenus varie en fonction de la concentration en aIF1-KA35 dans le milieu. Ce résultat est cohérent avec l'augmentation de Kd mesurée par anisotropie. Là encore, la concentration d'aIF1 nécessaire pour obtenir la saturation est très supérieure à celle prédite par la mesure du Kd en anisotropie. L'utilisation de l'ARNm très long pouvant aussi expliquer cet effet.

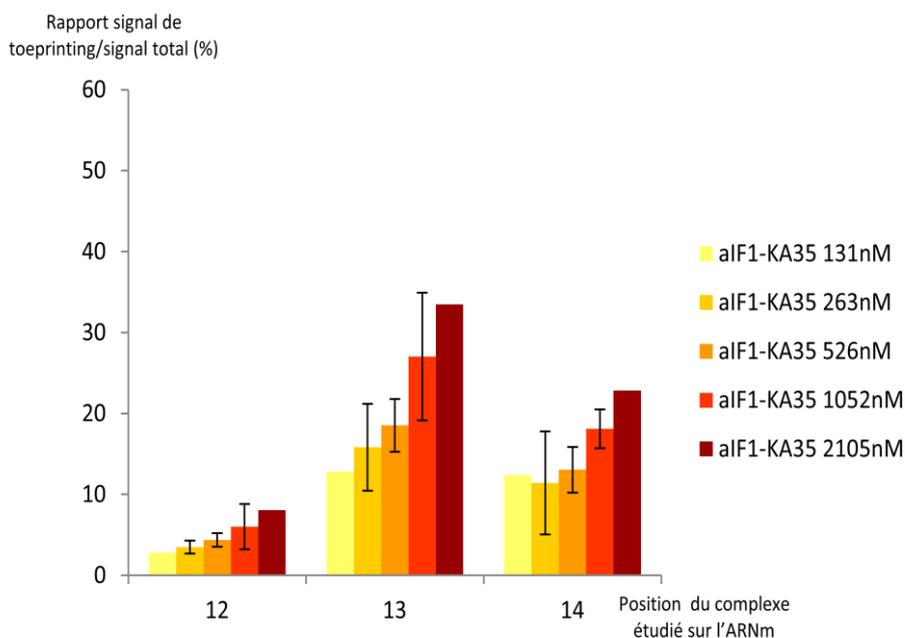


Figure 3.40 : Toeprint avec le complexe 30S:aIF1-KA<sub>35</sub> sur l'ARNm WT.  
Gamme de concentration en aIF1-KA<sub>35</sub>

*Les barres d'erreurs avec aIF1-KA<sub>35</sub> sont obtenus après 2 expériences indépendantes*

Par ailleurs, des expériences d'anisotropie de fluorescence avec le variant aIF1\*-KA35 titré avec la 30S seul ont révélé une très nette augmentation du K<sub>d</sub>, tout comme pour les expériences similaires réalisées avec aIF1\*. Le K<sub>d</sub> d'aIF1\*-KA35 pour le complexe 30S:ARNm est de 45nM et passe à 381 ± 46nM dans le cas d'un titrage avec la 30S seule. **Ces résultats nous indiquent que dans le cas du mutant aIF1\*-KA35, il y a toujours coopérativité entre la liaison d'aIF1 et la liaison de l'ARNm au ribosome. La lysine 35 est donc plutôt orientée vers la liaison de la 30S.**

L'étude de ce mutant dans la formation et la stabilité du CI devra être réalisée ultérieurement dans des conditions de concentration en aIF1-KA35 assurant la saturation de la 30S.

#### 3.2.4.4 Etude du double-mutant aIF1-RA32-YA33

Lors des expériences précédentes, nous avons pu mettre en évidence le rôle de l'arginine 32 et de la tyrosine 33 dans la stabilisation de la 30S sur l'ARNm, et dans la déstabilisation du complexe centré sur le codon de démarrage. La mutation conjointe de ces deux résidus pourrait renforcer encore les effets observés avec les mutants ponctuels. C'est pourquoi nous avons produit le double mutant aIF1-RA32-YA33.

### 3.2.4.4.1 Le complexe aIF1-RA32-YA33:30S sur l'ARNm WT : rôle dans la stabilisation de la 30S sur l'ARNm

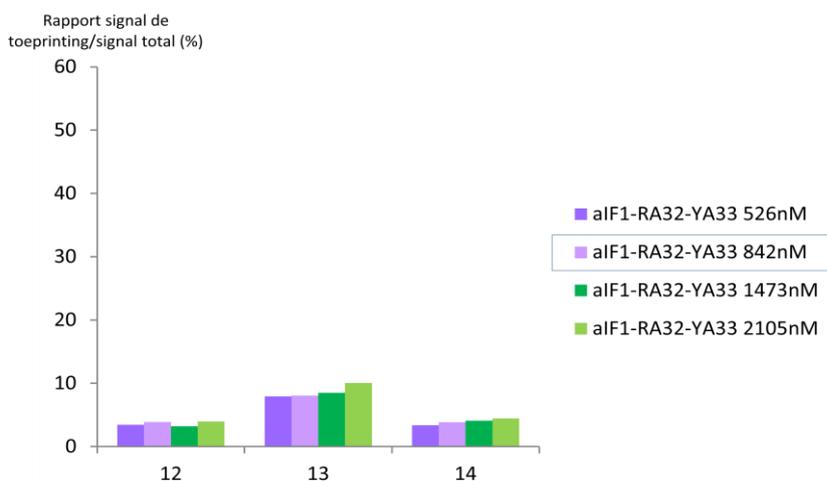


Figure 3.41 : Toeprint avec le complexe 30S:aIF1-RA<sub>32</sub>-YA<sub>33</sub> sur l'ARNm WT. Gamme de concentration en aIF1-RA<sub>32</sub>-YA<sub>33</sub>

Afin de contrôler l'effet de cette double mutation dans la stabilisation de la 30S sur l'ARNm, les expériences de toeprint ont tout d'abord été réalisées avec le complexe aIF1-RA32-YA33:30S et l'ARNmWT. Dans un premier temps, un Kd de 61 +/- 15nM de ce variant pour le complexe 30S ARNm a été déterminé par anisotropie de fluorescence. Afin de se placer dans des conditions de saturation de la 30S en aIF1-RA32-YA33, une gamme de concentration de 526 à 2105nM d'aIF1-RA32-YA33 a été réalisée (voir figure 3.41 ci-dessus). La suite des expériences sera réalisée à la concentration de 842nM de ce variant dans le milieu réactionnel.

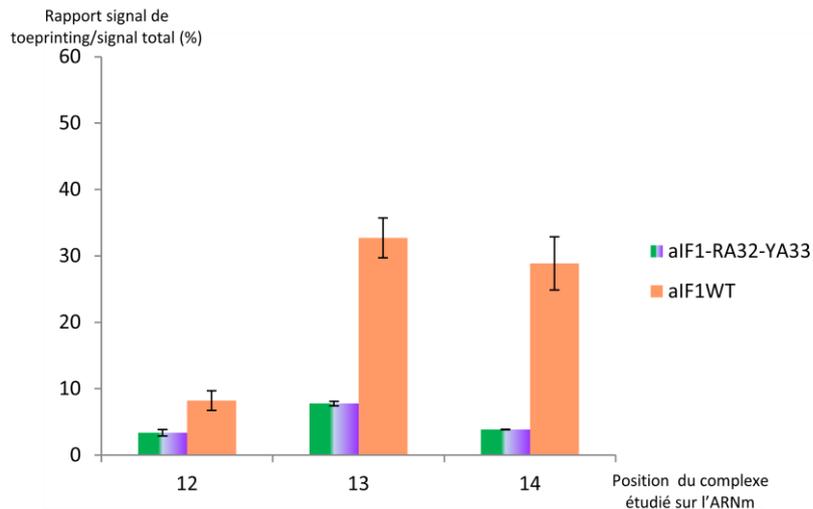


Figure 3.42 : Toeprint avec le complexe 30S:aF1-RA<sub>32</sub>-YA<sub>33</sub>, 30S:aF1WT sur l'ARNm WT.  
 Les barres d'erreurs avec aF1-RA<sub>32</sub>-YA<sub>33</sub> sont obtenues après 3 expériences indépendantes  
 Les barres d'erreurs avec aF1 WT sont obtenues avec plus de 6 expériences indépendantes

Les résultats du toeprint avec aF1-RA32-YA33:30S:ARNm sont montrés dans la figure 3.42 ci-dessus. Le double mutant ne possède plus qu'une faible capacité à stabiliser la 30S sur l'ARNm. Comme attendu, le profil obtenu est similaire à celui du mutant aF1-YA33 du fait de la présence de cette mutation dans ce variant. Un profil tout à fait similaire est obtenu en utilisant l'ARNm Val1Met2.

#### 3.2.4.4.2 Etude du variant aF1-RA32-YA33 dans le CI

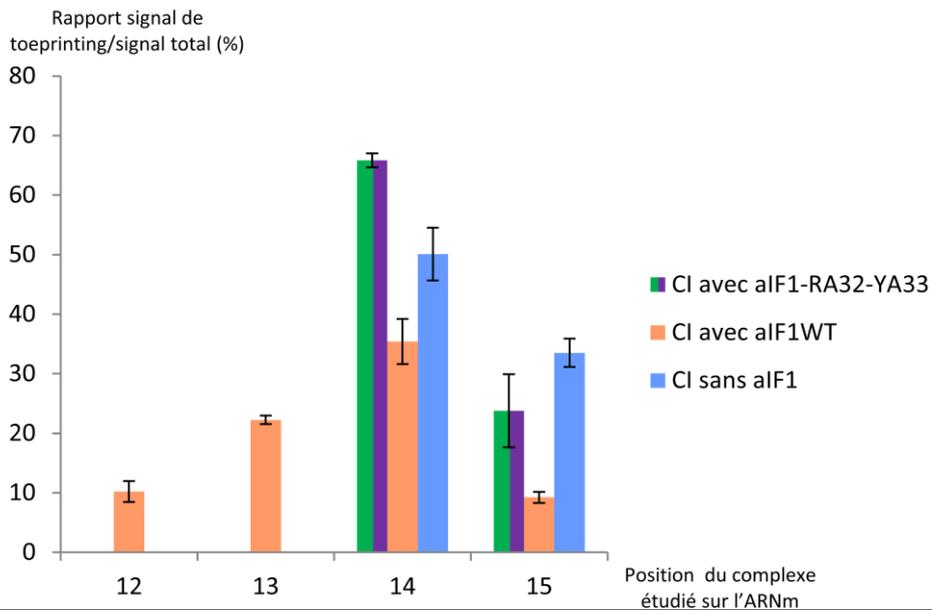


Figure 3.43 : Comparaison des profils de toeprint obtenus pour le CI en présence d'aF1WT ou d'aF1-RA<sub>32</sub>-YA<sub>33</sub> sur l'ARNm WT  
 Les barres d'erreurs avec aF1-RA<sub>32</sub>-YA<sub>33</sub> sont obtenues après 2 expériences indépendantes  
 Les barres d'erreurs avec aF1 WT et sans aF1 sont obtenues avec plus de 6 expériences indépendantes

Lorsque tous les facteurs et l'ARNt<sub>i</sub> (CI avec aIF1-RA32-YA33) sont présents dans le milieu réactionnel, le profil de toeprint est totalement décalé en position +14 et +15, correspondant aux positions attendues lorsqu'aIF1 n'est plus présent dans le milieu (voir figure 3.43 ci-dessus).

Le double mutant aIF1-RA32-YA33 n'est donc plus capable de déstabiliser le complexe centré sur le codon de démarrage, et cet effet est d'autant plus augmenté que ce variant a perdu en grande partie sa capacité de stabiliser le ribosome sur l'ARNm. De même, si l'ARNm Val1Met2 est utilisé lors de l'expérience de toeprint, on peut constater qu'en présence du CT avec aIF1-RA32-YA33, le profil obtenu est similaire au profil obtenu sans aIF1 dans le milieu réactionnel.

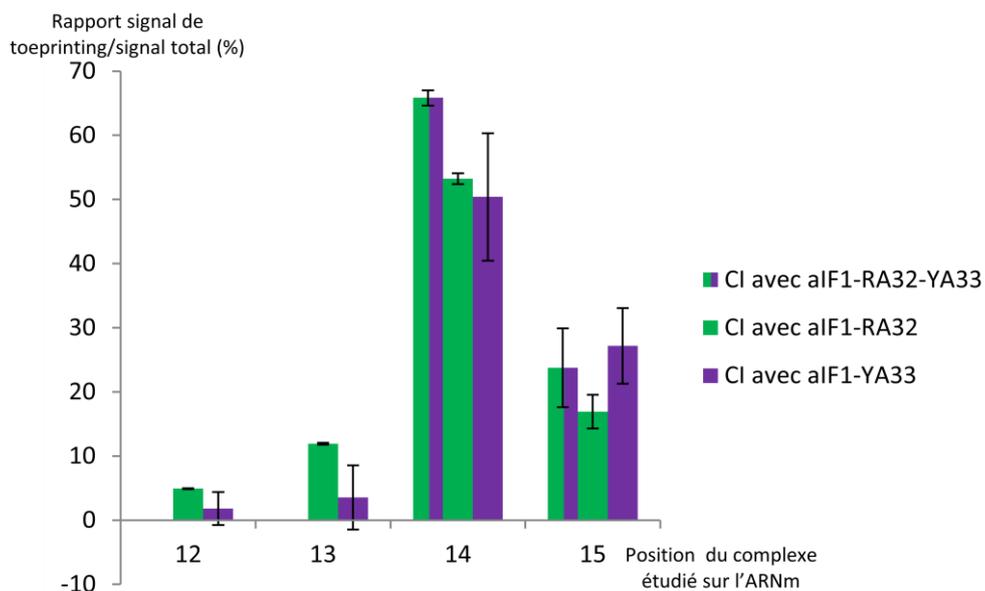


Figure 3.44 : Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1RA<sub>32</sub>, d'aIF1-YA<sub>33</sub> ou d'aIF1-RA<sub>32</sub>-YA<sub>33</sub> sur l'ARNm WT  
*Les barres d'erreurs sont obtenues après 2 expériences indépendantes*

De manière intéressante, si l'on compare ces résultats avec ceux obtenus avec les simples mutants aIF1-RA32 ou aIF1-YA33, on peut constater que le profil obtenu avec le double mutant se rapproche du profil obtenu avec aIF1-YA33. Toutefois, il n'y a plus du tout de signal en position +12 et +13 (voir figure 3.44 ci-dessus). Ces résultats semblent nous indiquer un effet cumulatif de ces mutations.

### 3.2.4.5 Etude du double-mutant aIF1-RA32-KA35

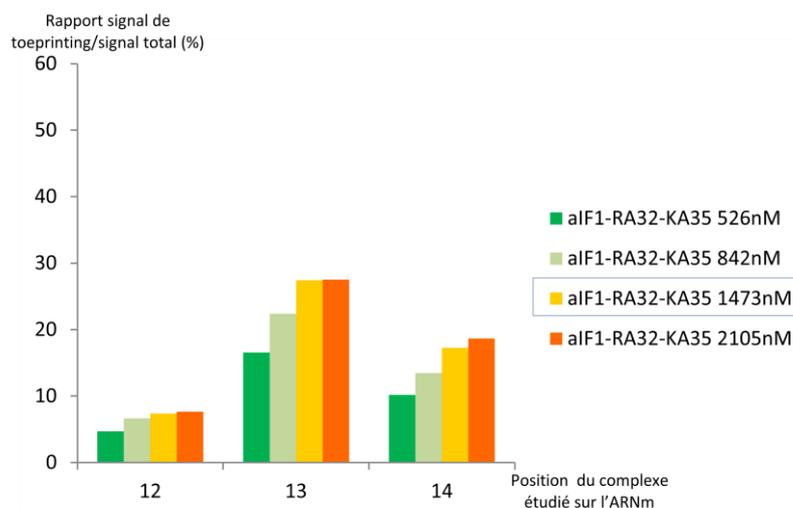


Figure 3.45 : Toeprint avec le complexe 30S:aIF1-RA<sub>32</sub>-KA<sub>35</sub> sur l'ARNm WT. Gamme de concentration en aIF1-RA<sub>32</sub>-KA<sub>35</sub>

Dans un premier temps, un Kd de  $167 \pm 10$  nM d'aIF1-RA32-KA35 pour le complexe 30S:ARNm a été déterminé par anisotropie de fluorescence. Ensuite, une gamme de concentrations a été réalisée dans des expériences de toeprint (voir figure 3.45 ci-contre). La suite des expériences se fera à la concentration de 1473nM de ce variant dans le milieu réactionnel.

#### 3.2.4.5.1 Le complexe aIF1-RA32-KA35:30S sur l'ARNm WT : rôle dans la stabilisation de la 30S sur l'ARNm

Les résultats du toeprint réalisé avec aIF1-RA32-KA35:30S:ARNm sont montrés dans la figure 3.46 ci-dessous. Comme attendu, la mutation conjointe des résidus R32 et K35 en alanine entraîne seulement une faible variation du profil d'empreinte. Par contre, la concentration saturante est seulement atteinte à 1473nM.

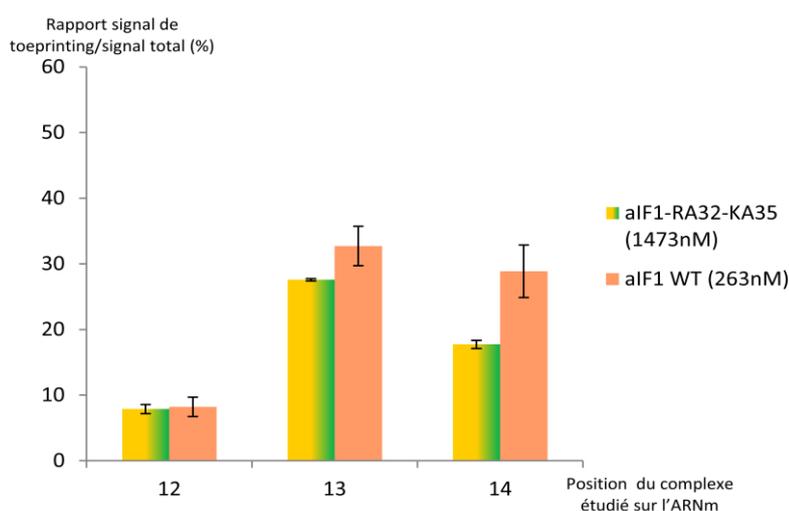


Figure 3.46: Toeprint avec le complexe 30S:aIF1-RA<sub>32</sub>-KA<sub>35</sub>, 30S:aIF1WT sur l'ARNm WT.

Les barres d'erreurs avec aIF1-RA<sub>32</sub>-KA<sub>35</sub> sont obtenues après 2 expériences indépendantes  
Les barres d'erreurs avec aIF1 WT sont obtenues avec plus de 6 expériences indépendantes

### 3.2.4.5.2 Etude du variant *aIF1-RA32-KA35* dans le CI

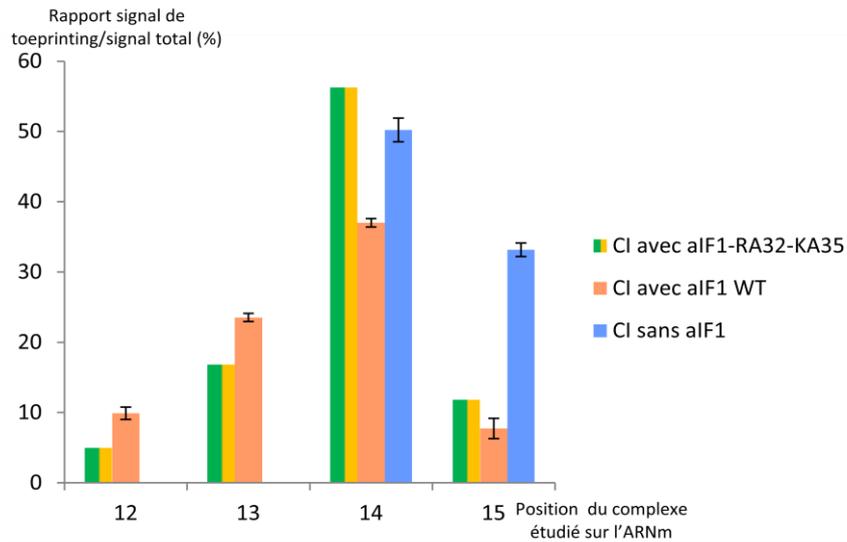


Figure 3.47 : Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1WT ou d'aIF1-RA<sub>32</sub>-KA<sub>35</sub> sur l'ARNm WT  
 Les barres d'erreurs avec aIF1 WT et sans aIF1 sont obtenus avec plus de 6 expériences indépendantes

On peut observer que le profil de toeprinting obtenu avec le CI (avec aIF1-RA32-KA35) est augmenté vers la position +14 (voir figure 3.47 ci-dessus).

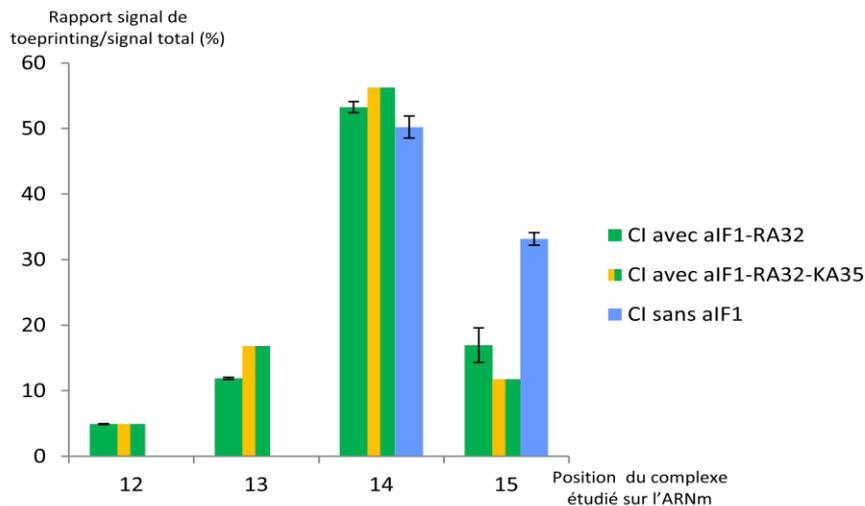


Figure 3.48 : Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1RA<sub>32</sub>, d'aIF1-KA<sub>35</sub> ou d'aIF1-RA<sub>32</sub>-KA<sub>35</sub> sur l'ARNm WT  
 Les barres d'erreurs sont obtenus après au moins 2 expériences indépendantes

De plus, lorsque l'on compare les profils d'empreintes obtenus avec le mutant ponctuel aIF1-RA32, on constate qu'ils sont très similaires. Le profil est décalé vers les positions +14 et +15. Cependant, les positions +12 et +13 restent présentes (figure 3.48 ci-dessus). Ces résultats nous indiquent

qu'aIF1-RA32-KA35 conserve tout de même une petite partie de sa fonction déstabilisatrice du complexe de démarrage centré sur le codon AUG.

### 3.2.5 Etude du rôle du domaine N-terminal d'aIF1 dans la formation du complexe de démarrage de la traduction.

Comme nous avons pu le constater dans les expériences d'anisotropie de fluorescence, l'extrémité N-terminale flexible d'aIF1 (résidus 1-23) n'a pas de fonction dans la liaison d'aIF1 à la sous-unité du ribosome ( $K_d=0.7 \pm 0.1\text{nM}$ ). Cependant, cette extrémité pourrait tout de même jouer un rôle dans la fonction d'aIF1 lors du démarrage de la traduction. C'est pourquoi nous avons étudié le variant d'aIF1 délété de son extrémité N-terminale (aIF1- $\Delta$ N) par toeprint. Comme attendu d'après le  $K_d$  d'aIF1\* -  $\Delta$ N pour le complexe 30S:ARNm, aucune variation d'intensité de toeprint n'est observée lorsqu'une gamme de concentration de 263nM à 2105nM d'aIF1- $\Delta$ N dans le milieu réactionnel est réalisée. La suite des expériences se fera en utilisant une concentration d'aIF1- $\Delta$ N de 263nM dans le test.

#### 3.2.5.1 Etude du complexe aIF1- $\Delta$ N:30S:ARNm

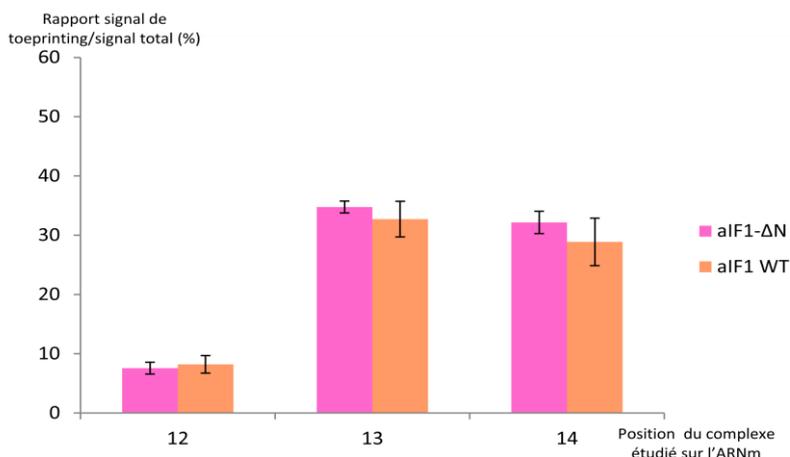


Figure 3.49 : Toeprint avec le complexe 30S:aIF1- $\Delta$ N, 30S:aIF1WT et l'ARNm WT.

Les barres d'erreurs avec aIF1- $\Delta$ N sont obtenues après 3 expériences indépendantes  
Les barres d'erreurs avec aIF1 WT sont obtenues avec plus de 6 expériences indépendantes

Les résultats obtenus avec le complexe aIF1- $\Delta$ N:30S:ARNm sont montrés dans la figure 3.49 ci-contre. On peut constater que le profil et l'intensité des empreintes obtenus est très similaire à ceux obtenus avec aIF1WT. L'extrémité N-terminale d'aIF1 n'est pas impliquée dans la stabilisation de la 30S à l'ARNm. Ceci peut s'expliquer par la présence attendue de cette extrémité du côté opposé à la 30S.

### 3.2.5.2 Etude de l'extrémité N-terminale dans le CI

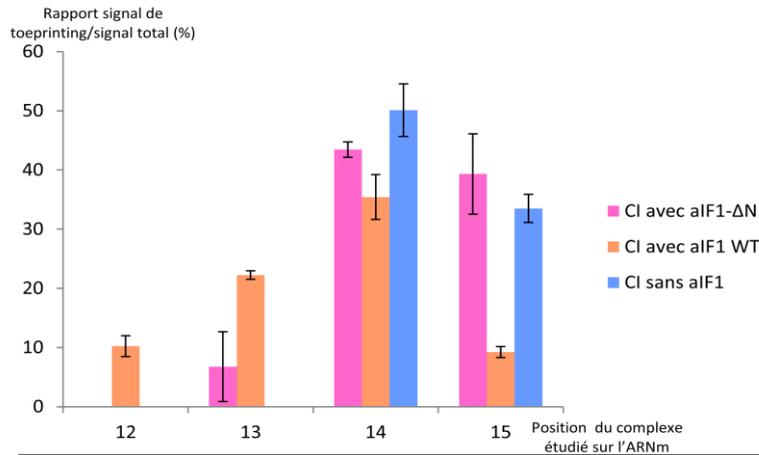


Figure 3.50 : Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1WT ou d'aIF1-ΔN sur l'ARNm WT  
 Les barres d'erreurs avec aIF1 WT et sans aIF1 sont obtenues avec plus de 6 expériences indépendantes  
 Les barres d'erreurs avec aIF1-ΔN sont obtenues avec 3 expériences indépendantes

Lorsque les expériences de toeprint avec le complexe total et aIF1-ΔN sont réalisées avec l'ARNm WT, on constate que le profil et l'intensité de toeprinting obtenu est très proche de ceux obtenus lorsqu'aIF1 n'est plus présent dans le milieu réactionnel (voir figure 3.50 ci-dessus). **De façon intéressante, l'extrémité N-terminale d'aIF1 joue un rôle dans la déstabilisation du complexe centré sur le codon de démarrage.** En effet, sa présence dans le complexe de démarrage pourrait entraîner une conformation particulière du ribosome sur l'ARNm et ainsi faciliter la déstabilisation du complexe au niveau du codon AUG lors du démarrage de la traduction.

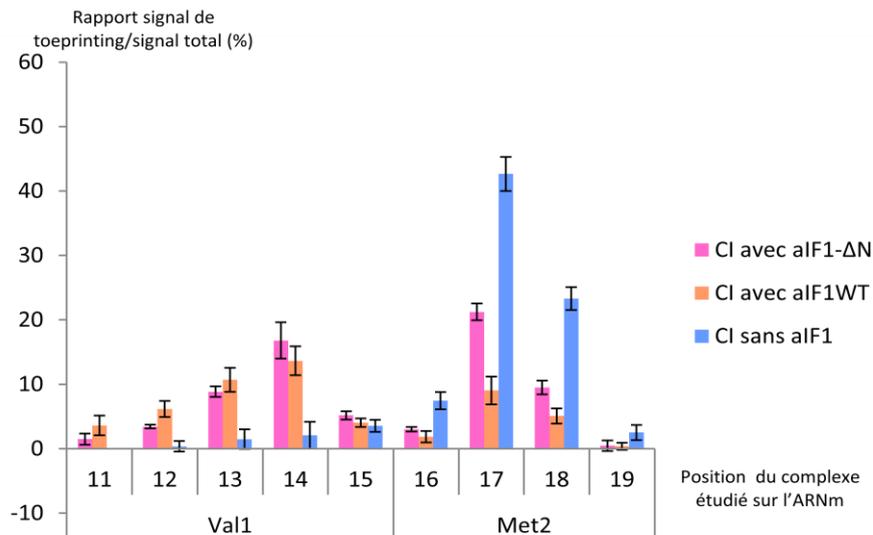


Figure 3.51 : Comparaison des profils de toeprint obtenus pour le CI en présence d'aIF1WT ou d'aIF1-ΔN sur l'ARNm Val1Met2  
 Les barres d'erreurs avec aIF1 WT et sans aIF1 sont obtenues avec plus de 6 expériences indépendantes  
 Les barres d'erreurs avec aIF1-ΔN sont obtenues avec 4 expériences indépendantes

Lorsque l'ARNm Val1Met2 est utilisé, le profil de toeprint est augmenté en position +17 et +18 (voir figure 3.51 ci-contre), indiquant aussi une perte de la fonction déstabilisatrice du complexe centré sur le codon de démarrage d'aIF1 lorsque l'extrémité N-terminale est absente. Malgré tout, l'effet est moins marqué lorsque le codon méthionine est en position 2, et donc dans un contexte moins favorable à la formation du CI.

### **3.2.6 VI : Rôle d'aIF1 dans le rejet d'un mauvais complexe de démarrage**

#### **3.2.6.1 Utilisation d'un ARNt élongateur dans la formation du complexe de démarrage de la traduction.**

Nous avons pu constater que la présence de l'ARNti était essentielle à la formation du complexe de démarrage centré sur le codon AUG de l'ARNm. De manière surprenante, l'étude de la spécificité d'aIF2 de *Pyrococcus abyssi* vis-à-vis des ARNt montre que l'hétérotrimère n'est pas capable de distinguer efficacement un ARNt méthionylé de type élongateur d'un ARNt méthionylé de type initiateur. En effet, aIF2 de *P. abyssi* lie le Met-ARNt<sub>i</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> avec un Kd de 90nM et le Met-ARNt<sub>m</sub><sup>Met</sup> (élongateur) avec un Kd de 450nM. Le taux de discrimination de l'ARNt initiateur par rapport à l'ARNt élongateur n'est donc qu'un facteur 5.

Ainsi, on peut se demander comment un éventuel complexe aIF2:Met-ARNt<sub>m</sub><sup>Met</sup> est exclu du démarrage de la traduction. Pour répondre à cette question, nous avons étudié l'influence de la présence d'un ARNt élongateur méthionine (Met-ARNt<sub>m</sub><sup>Met</sup>) dans le CI. Tout d'abord, une gamme de concentration d'ARNt élongateur méthionylé a été réalisée dans les mêmes conditions qu'avec l'ARNt initiateur (voir résultats précédents). En se plaçant à une concentration de 1052nM de Met-ARNt<sub>m</sub><sup>Met</sup>, nous nous assurons d'être en condition saturante d'aIF2 et de 30S.

Dans une première série d'expériences, nous avons regardé la stabilité des complexes obtenus avec l'ARNt initiateur ou avec l'ARNt élongateur en absence d'aIF1.

On peut constater dans la figure 3.52 ci-dessous que lorsque le facteur aIF1 n'est pas présent dans le CI, l'empreinte du ribosome sur l'ARNm est similaire en présence d'un ARNt initiateur ou d'un ARNt élongateur méthionylé.

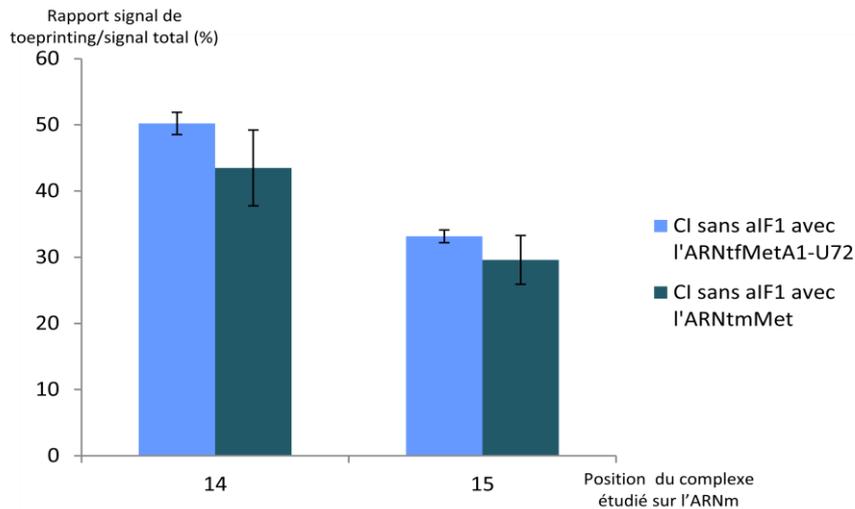


Figure 3.52 : Comparaison des profils de toeprint obtenus avec le CI sans aIF1 en présence de l'ARNt<sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> ou de l'ARNt<sup>Met</sup>  
 Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes

Au contraire, lorsque le facteur aIF1 est ajouté au milieu réactionnel, le profil de toeprint obtenu avec l'ARNt élongateur méthionylé est plus fortement diminué en position +14 par rapport à celui obtenu avec un ARNt initiateur (voir figure 3.53 ci-dessous). Le complexe de démarrage est donc plus déstabilisé de la position centrée sur le codon AUG. Ainsi, en absence d'aIF1, le complexe de démarrage se forme sur l'ARNm de façon indépendante de la présence d'un ARNt élongateur méthionylé ou initiateur dans le milieu. Au contraire, lorsque le facteur aIF1 est ajouté au milieu réactionnel, l'action déstabilisatrice d'aIF1 du complexe de démarrage centré sur le codon AUG est augmentée pour le complexe formé avec l'ARNt élongateur. Ainsi, aIF1 semble jouer un rôle dans la discrimination des ARNt initiateurs et élongateur en favorisant la formation des complexes possédant un ARNt initiateur au niveau du site P du ribosome.

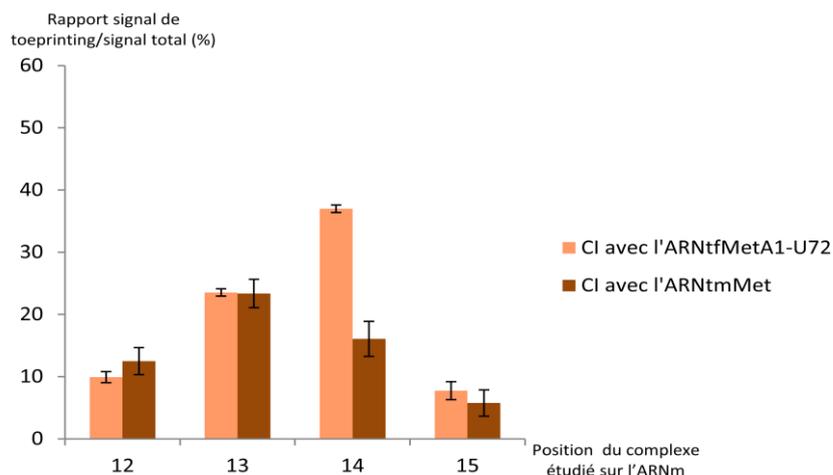


Figure 3.53 : Comparaison des profils de toeprint obtenus pour le CI en présence de l'ARNt<sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> ou de l'ARNt<sup>Met</sup>  
 Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes

### 3.2.6.2 Etude de l'influence des 3 paires de bases GC de l'ARN<sub>t</sub><sub>i</sub> dans la formation du complexe de démarrage de la traduction sur l'ARN<sub>m</sub>.

Comme cela est décrit dans le chapitre 1, les ARN<sub>t</sub><sub>i</sub> possèdent trois paires de bases GC caractéristiques dans leur tige anticodon, les distinguant en particulier des ARN<sub>t</sub> élongateurs. Afin d'étudier l'importance de ces bases dans la discrimination par aIF1 dans le CI des ARN<sub>t</sub> initiateurs et élongateurs, les mêmes expériences de toeprinting que précédemment ont été réalisées avec un variant de l'ARN<sub>t</sub><sub>m</sub><sup>Met</sup> possédant trois paires de bases GC dans sa tige anticodon (ARN<sub>t</sub><sub>m</sub><sup>Met</sup>3G3C).

Comme lors de l'expérience précédente, nous pouvons constater que lorsqu'aIF1 n'est pas présent dans le CI, les empreintes de toeprinting obtenus sont similaires, ne tenant pas compte de l'espèce d'ARN<sub>t</sub> présent dans le milieu (figure 3.54 ci-dessous).

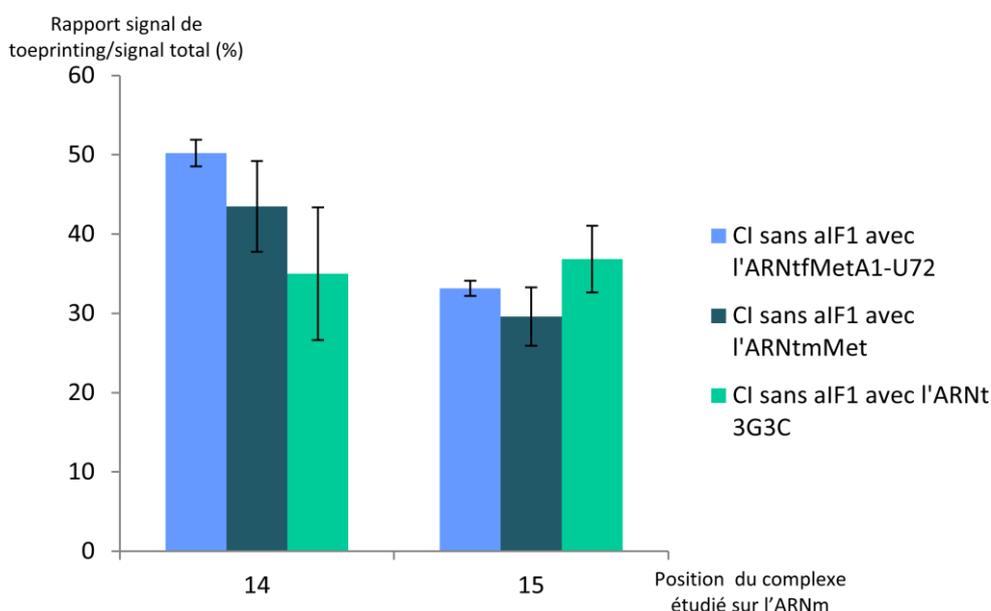


Figure 3.54 : Comparaison des profils de toeprinting obtenus pour le CI sans aIF1 en présence de différents ARN<sub>t</sub>

*Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes*

Là encore, lorsqu'aIF1 est absent, l'appariement codon:anticodon assure la stabilisation du complexe de démarrage.

Lorsque le facteur aIF1 est ajouté au CI contenant l'ARN<sub>t</sub><sub>m</sub><sup>Met</sup>3G3C, le profil de toeprinting obtenu redevient similaire à celui obtenu lorsque l'ARN<sub>t</sub><sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> est présent dans le CI (voir figure 3.55 ci-dessous). La différence entre l'ARN<sub>t</sub><sub>m</sub><sup>Met</sup>3G3C et l'ARN<sub>t</sub><sub>m</sub><sup>Met</sup> réside dans la présence des trois paires de bases GC dans la tige anticodon. Nous pouvons en conclure que la présence de ces trois bases GC dans la tige anticodon de l'ARN<sub>t</sub> méthionylé élongateur permet de rétablir un profil similaire à celui obtenu avec un ARN<sub>t</sub> initiateur.

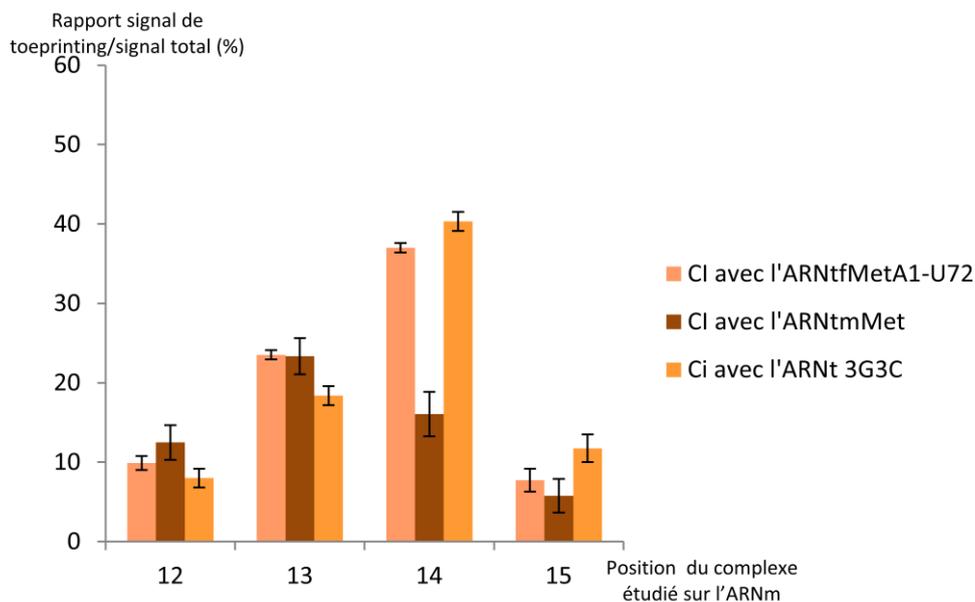


Figure 3.55: Comparaison des profils de toeprint obtenus pour le CI en présence de différents ARNt

*Les barres d'erreurs sont obtenues après plus de 6 expériences indépendantes*

De plus, la discrimination des ARNt initiateur et élongateur ainsi que le rétablissement du profil « initiateur » avec l'ARNt 3G3C n'est réalisée qu'en présence du facteur aIF1 dans le CI. La discrimination des ARNt initiateurs et élongateurs pourraient alors être réalisée par le facteur aIF1 via les trois paires de bases GC présentes dans la tige anticodon des ARNt initiateurs. Une explication pourrait être la présence du conflit stérique entre la boucle basique d'aIF1 et l'ARNt initiateur au niveau de la tige anticodon. Ces trois paires de bases conférant à l'ARNt initiateur une plus grande flexibilité qu'à un ARNt élongateur, le CI centré sur le codon AUG serait moins déstabilisé par aIF1 que lorsqu'un ARNt élongateur est présent dans le milieu. Cette caractéristique favoriserait alors la formation de « bons » complexes de démarrages centrés sur le codon AUG de l'ARNm.

**Ces expériences montrent que le facteur aIF1 déstabilise davantage un mauvais complexe de démarrage, formé avec un ARNtmMet (élongateur) ne possédant pas les trois paires de bases GC. Aussi, les trois paires GC sont importantes pour la sélection du bon complexe de démarrage.**

**Finalement, cette expérience renforce l'idée que la fonction d'aIF1 dans la déstabilisation du complexe de démarrage centré sur le codon AUG est associée à une activité d'édition des mauvais complexes de démarrage.**



# **4 Etude structurale du complexe de démarrage de la traduction archée**

Au début de ma thèse, très peu de données structurales étaient mise à notre disposition concernant les facteurs de démarrage de la traduction archée et aucun sur le complexe de démarrage. Ainsi, le laboratoire s'est attaché à résoudre la structure du complexe de démarrage archéen. Ces travaux se sont inscrits dans une dynamique globale du laboratoire consistant à obtenir les différents acteurs du démarrage archée afin d'étudier la structure du complexe par Cryo-EM. Ma contribution à ce travail a été d'obtenir les différents acteurs du démarrage permettant la formation du complexe de démarrage.

# Initiator tRNA accommodation during translation initiation mRNA probing

Pierre-Damien Coureux<sup>1</sup>, Christine Lazenec-Schurdevin<sup>1</sup>, Auriane Monestier<sup>1</sup>, Eric Larquet<sup>2</sup>, Lionel Cladière<sup>1&</sup>, Bruno P Klaholz<sup>3</sup>, Emmanuelle Schmitt<sup>1§</sup> and Yves Mechulam<sup>1§</sup>

<sup>1</sup>Laboratoire de Biochimie, Ecole polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau cedex, France

<sup>2</sup>Laboratoire de Physique de la Matière Condensée, Ecole polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau cedex, France

<sup>3</sup>Department of Structural Biology and Genomics, Institute of Genetics and Molecular and Cellular Biology, Centre National de la Recherche Scientifique (CNRS)/ Institut National de la Santé et de la Recherche Médicale (INSERM)/ Université Louis Pasteur, BP 10142, 67404 Illkirch, France

<sup>&</sup>Present address : Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France

<sup>§</sup>Co-senior author

Corresponding authors:

Pierre-Damien Coureux, E-mail: [pierre-damien.coureux@polytechnique.edu](mailto:pierre-damien.coureux@polytechnique.edu)

Emmanuelle Schmitt, E-mail: [emmanuelle.schmitt@polytechnique.edu](mailto:emmanuelle.schmitt@polytechnique.edu)

**Running title:** structure of translation initiation complexes

**En révision**

**Keywords:** translation initiation, tRNA, ribosome, cryo-EM.

**Eukaryotic and archaeal translation initiation complexes have a common structural core comprising e/aIF1, e/aIF1A, the ternary complex (TC, e/aIF2-GTP-Met-tRNA<sub>i</sub><sup>Met</sup>) and mRNA bound to the small ribosomal subunit. e/aIF2 plays a crucial role in this process but how this factor controls start codon selection remains unclear. Here, we present cryo-EM structures of the full archaeal 30S initiation complex showing two unprecedented conformational states of the TC. In the first state, the TC is bound to the ribosome in a relaxed conformation with the tRNA oriented out of the P site. In the second state, the tRNA is accommodated within the peptidyl (P) site and the TC becomes constrained. This constraint is compensated by codon/anti-codon base pairing whereas in the absence of a start codon, aIF2 contributes to swing out the tRNA. This spring force concept highlights a mechanism of codon/anticodon probing by the initiator tRNA directly assisted by aIF2.**

In all living cells, translation initiation allows accurate selection of the initiation codon on a messenger RNA, which then defines the reading frame of the protein to be synthesized. In eukaryotes, this process is highly complicated and is the target of many regulations<sup>1-3</sup>. The mechanism begins with the assembly of the ternary complex (TC), consisting of initiator methionyl-tRNA (Met-tRNA<sub>i</sub>) anchored to the GTP-bound form of the heterotrimeric initiation factor 2 (eIF2). The TC then binds to the small ribosomal subunit (40S) in the presence of two small initiation factors eIF1 and eIF1A. TC binding is assisted by eIF5<sup>4</sup> and by the multimeric factor eIF3<sup>5</sup>. eIF5 favors reversible GTP hydrolysis on eIF2. In the presence of factors belonging to the eIF4 family, the pre-initiation ribosomal complex (PIC) is recruited near the 5'-capped end of the mRNA and scans the mRNA to search for an AUG codon in an appropriate context. AUG recognition stops scanning, leads to factors release and to the assembly of an elongation-proficient 80S complex through large subunit joining, with the help of eIF5B. A large amount of data from genetics and biochemistry has been collected in order to decipher the molecular steps ensuring the correct selection of the AUG codon

on a messenger RNA<sup>6-10</sup>. In particular, isolation of yeast mutants able to initiate translation on non-AUG codon highlighted the key roles of eIF1 and eIF2 in the selection of the start codon. In the current model, AUG recognition triggers a conformational change of the 40S subunit, which leads to eIF1 departure, and then to the release of inorganic phosphate (Pi) from eIF2. In its GDP-bound form, eIF2 loses its affinity for the initiator tRNA and therefore leaves the 40S subunit<sup>11,12</sup>. Recent structural descriptions of eukaryotic translation initiation complexes have highlighted some of the structural changes that lead to a closed state with the initiator tRNA base-paired to the start codon<sup>13-16</sup>. These structures have shown how accommodation of the initiator tRNA at the P site, from a P<sub>OUT</sub> to a P<sub>IN</sub> position<sup>7,17</sup>, is stabilized by eIF1A and is accompanied by a closure of the 40S subunit as well as by the destabilization of eIF1 binding. However, these structures do not give any clue on how the departure of eIF1 triggers the release of the Pi group from eIF2 and thereby the release of eIF2-GDP. Moreover, the absence of contact between the central GTP-binding  $\gamma$  subunit of eIF2 and the ribosome remains intriguing according to radical probing experiments that suggested interaction of domain III of eIF2 $\gamma$  with h44 in rRNA<sup>18</sup>.

In archaea, long-range scanning does not occur due to the occurrence of Shine-Dalgarno sequences or of very short 5'UTR on mRNA. However, archaeal translation initiation uses three initiation factors, aIF1, aIF1A and aIF2, homologous to their eukaryotic counterparts. This suggests that the archaeal factors fulfill similar functions as the eukaryotic factors<sup>19</sup>. Such a parallel relies on the idea that the detection of the start codon in archaea proceeds through a local scanning mechanism within a short mRNA region compatible with maintenance of SD:antiSD pairing. This leads to the definition of a structural core, common to eukaryotes and archaea, that controls translation initiation accuracy. This core is made of the small ribosomal subunit, the mRNA, the Met-tRNA<sub>i</sub><sup>Met</sup> and the three factors e/aIF1, e/aIF1A and e/aIF2.

To further address the role of e/aIF2, we have analyzed the full archaeal 30S initiation complex (30S IC). We have determined two cryo-EM structures revealing the existence of two

unprecedented conformational states of the 30S-bound TC. In the first complex the tRNA is oriented out of the peptidyl (P) site in a standby position (the complex was therefore called IC0-P<sub>OUT</sub>) and the heterotrimeric aIF2 strongly interacts with the 30S subunit and with aIF1. In the second complex, the initiator tRNA is accommodated within the P site (the complex was therefore called IC1-P<sub>IN</sub>) and partially releases interactions with aIF2, reflecting the structural transition from the standby position to the decoding position. Comparison of IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub> shows that the TC is constrained when the tRNA is engaged in the P site. However, upon start codon recognition this constraint is compensated by codon/anti-codon base pairing. In the absence of a start codon, aIF2 contributes to swing out the tRNA from the P site. This spring force concept provides a novel view of the role of e/aIF2 in start codon selection. Moreover, we observed contacts between aIF1 and the aIF2 $\gamma$  switch regions known to control its nucleotidic state. This gives a structural framework to explain how aIF1 departure provokes the release of Pi from aIF2. According to previous biochemical and genetic studies in eukaryotic systems<sup>7,9,18</sup>, the features of the two novel states of the core initiation complex are likely to be relevant for eukaryotic translation initiation.

## RESULTS

### Assembly and structural analysis of the full archaeal initiation complex

We assembled the full initiation complex using archaeal 30S subunits from *Pyrococcus abyssi* (Pa-30S), initiation factor Pa-aIF1, Pa-aIF1A, the ternary complex (Met-tRNA<sub>i</sub><sup>Met</sup> A1-U72 variant from *E. coli* (Supplementary Fig. 1a), GDPNP and Pa-aIF2 composed of 3 subunits  $\alpha$ ,  $\beta$  and  $\gamma$ ) and a synthetic 26 nucleotide long mRNA corresponding to the natural start region of Pa-aEF1A, that contains a strong Shine-Dalgarno sequence (A<sub>(-17)</sub>UUUGGAGGUGAUUAAA<sub>(+1)</sub>UGCCAAAG<sub>(+9)</sub>). The complex was purified by affinity chromatography using N-terminally tagged versions of Pa-aIF1, Pa-aIF2 $\alpha$  and Pa-aIF2 $\beta$  (Supplementary Fig. 1b). An excess of mRNA, aIF1A and TC was added before spotting onto the grids. Cryo-EM images were collected on an FEI Cs-corrected Titan Krios

microscope operated at 300kV (Supplementary Fig. 1c). After image processing and 3D classification using RELION<sup>20</sup>, the structures of two complexes were identified (Fig. 1 and Supplementary Fig. 1d). In one complex, the tRNA is totally out of the P site and is therefore referred to as IC0-P<sub>OUT</sub>. The other structure, featuring a tRNA accommodated in the P site was named IC1-P<sub>IN</sub>. The two electron density maps were obtained at overall resolutions of 5.3 and 7.5 Å for IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub>, respectively, according to the 0.143 FSC criterion<sup>21</sup> (Supplementary Fig. 1e). As already observed for other initiation complexes<sup>22</sup>, local resolution is best in the core of the 30S subunit and decreases as distance from the 30S increases (Supplementary Fig. 1g and h). Model construction was mainly based on the placement of known structures of archaeal homologous components (Supplementary Fig. 2 and Online methods). In both IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub>, consistently with the previous structures of eukaryotic complexes, aIF1A and aIF1 were present in the A site and in front of the P site, respectively<sup>15,16,22,23</sup>. The mRNA was clearly visible in the electron density. Interaction of the SD sequence with the anti-SD was modeled and the start codon was positioned within the P site (Supplementary Fig. 3). mRNA was not visible from base +5 to +9 at the 3' extremity, as in yeast 48S complex (py48S)<sup>14</sup>; bases -17 and -16 at the 5' end were not visible.

### **The IC0-P<sub>OUT</sub> structure highlights a new network of interactions around aIF2**

In IC0-P<sub>OUT</sub>, all actors of translation initiation are bound to the ribosome but the initiator tRNA is located totally outside of the P site (Figs. 1a and 2a). The atomic model corresponding to the archaeal TC<sup>24</sup> could be readily fitted in the electron density. Therefore, on the 30S subunit, the initiator tRNA is bound to aIF2 in a manner similar to that in the free TC, without constraints induced by the ribosome. As compared to the crystal structure, only slight rotations of the rigid unit made of domains 1 and 2 of aIF2 $\alpha$  (aIF2 $\alpha$ -D12) were observed, accompanied by a minor movement of the tRNA (Supplementary Fig. 2b).

Within the heterotrimeric aIF2, the  $\gamma$  subunit is tightly bound to the 30S subunit, mainly held by two contact points (Figs. 2 and 3). One striking interaction involves domain III of aIF2 $\gamma$  (aIF2 $\gamma$ -DIII). The C-terminal part (residues 351 to 354) of the long loop between  $\beta$ 16 and  $\beta$ 17 as well as residues 372 to 375 from the loop between  $\beta$ 18 and  $\beta$ 19 overhang the top of the h44 bulge comprising bases G<sub>1377</sub>, A<sub>1378</sub> and bases G<sub>1437</sub>, A<sub>1438</sub>. Accordingly, a distortion of the helical structure of h44 is observed and the local structure of h44 had to be remodeled (Fig. 3c). Moreover, the C-terminus of ribosomal protein uL41 points toward the  $\beta$ 17 side of aIF2 $\gamma$ -DIII (Fig. 3). A second contact involves the long  $\beta$ -hairpin ( $\beta$ 10- $\beta$ 11) of aIF2 $\gamma$ -DII and the groove formed by the rRNA h27 loop on one side (856-860) and by the 3' extremity of h24 (723-725) on the other side. Notably, this region of the rRNA links the decoding center and the platform. This contact gives sense to the presence of the  $\beta$ 10- $\beta$ 11 hairpin in all e/aIF2 $\gamma$ , whereas it is absent in elongation factors e/aEF1A<sup>25</sup>. The position of aIF2 $\alpha$ -D3 with respect to aIF2 $\gamma$  corresponds to that observed in the TC and in other aIF2 structures<sup>26-28</sup>. aIF2 $\alpha$ -D3 is sitting on the platform close to the tip of h23 and h24 bulges. Moreover, the C-terminus of aIF2 $\alpha$  points towards h24 (Fig. 3b). Interestingly, the length of the e/aIF2 $\alpha$  C-terminal tail varies from species to species. In eukaryotes, this tail was shown to negatively influence tRNA binding, and it was suspected that this negative effect might be cancelled upon interaction of the TC with the ribosome<sup>24,29</sup>. In this context, it is notable that a contact between the C-terminal tail of eIF2 $\alpha$  and h24 can be inferred from the position of aIF2- $\alpha$ D3 observed here. Release of such a contact may help the release of eIF2 from the tRNA upon start codon recognition.

On the side of aIF2 $\alpha$ -D12, electron density suggests contacts with ribosomal protein eS1. In the case of py48S eIF2 $\alpha$ -D12 was located in the E site, close to the -2 and -3 positions of mRNA. This position is suitable for a participation of eIF2 $\alpha$  in sensing the favorable environment of the correct start codon<sup>30</sup>. The different orientation of aIF2 $\alpha$ -D12 observed here is possibly related to the pre-positioning of the 30S subunit on the mRNA due to the recognition of the SD sequence.

Finally, the electron density also suggests interaction of the  $\alpha\beta$  domain of aIF2 $\beta$  with the D-

stem and D-loop of the tRNA (Figs. 1a and 3d). Such a position of aIF2 $\beta$  with respect to the tRNA and to aIF2 $\gamma$  is similar to that observed in eukaryotic translation initiation complexes<sup>14</sup>. Hence, on the whole, the initiator tRNA appears to be wrapped up by the three aIF2 subunits leaving the anticodon loop free of interaction (Figs. 1a and 3d). Notably, the observed position of the ternary complex on the 30S subunit prevents uncontrolled tRNA access to the ribosomal E site.

In front of the aIF2 $\gamma$ -DIII domain, aIF1 occupies its previously described position on h44 close to the P site<sup>15,16,22</sup> (Fig. 1a). As compared to the structure of the isolated factor, the conformation of the basic L1 loop had to be remodeled to account for the electron density. The loop comes close to the P site codon on the mRNA, suggesting that aIF1 contributes to the stabilization of the mRNA codon within the P site. A key feature of the present structure is that aIF1 and aIF2 interact with each other (Fig. 2a). One contact between both initiation factors involves the  $\beta$ 20- $\beta$ 21 loop of aIF2 $\gamma$ -DIII (402-404) with the top of the aIF1  $\alpha$ 2 helix and the following loop 91-95. A supplementary density at the N-terminus of aIF1 corresponds to a part of the 23 N-terminal residues missing in aIF1 crystal structure, and shows the occurrence of a contact between the N-terminal domain of aIF1 and the two switch regions of aIF2 $\gamma$ . Finally, the electron density of aIF1 in IC0-P<sub>OUT</sub> is as strong as that observed for the ribosome, reflecting its tight binding (Fig. 1a and Supplementary Fig. 1g).

aIF1A is positioned within the A site as in eukaryotic PIC<sup>15,16,22</sup>. Interactions between h44 and the tight turn 34-36 are visible. To avoid clashes with aIF1A, bases 1447 and 1448 must flip out from h44, as previously observed<sup>15,16,22</sup>. Notably, the channel leading to the P site, occupied by the functionally important N-terminal extremity of eIF1A<sup>31</sup> in the py48S complex<sup>22</sup>, is empty in the present structure (Supplementary Fig. 3b). The N-terminal tail of Pa-aIF1A is 22-residue long, ten residues shorter than in the eukaryotic factor. Residues 13 to 22 form an elongated peptide with its extremity facing the h32-h34 junction in the 30S head but the first 12 residues are not visible within the electron density.

### IC1-P<sub>IN</sub> shows a constrained state of the TC

In IC1-P<sub>IN</sub>, the initiator tRNA is positioned in the P site (Fig. 1b and 2b). The density is consistent with the interaction of the anticodon with the start codon on mRNA whereas the 3' methionylated extremity remains bound to aIF2 $\gamma$ . The structure of the P site initiator tRNA as observed in the *T. thermophilus* 70S complex<sup>32</sup> could not be readily rigid-body fitted in the density. Indeed a bending of the tRNA molecule at the level of the hinge region was necessary (Fig 4c and Supplementary Fig. 4c). The acceptor end was also remodeled to fit the 3' methionyl-adenosine in its pocket on aIF2 $\gamma$ -DII. Notably, aIF2 $\gamma$  moves only very slightly in IC1-P<sub>IN</sub> as compared to IC0-P<sub>OUT</sub> and remains bound to h44 via its domain III (Figs. 2b and 5a). Hence, in IC1-P<sub>IN</sub>, the tRNA molecule interacts with aIF2 in a manner different from that in free TC or in IC0-P<sub>OUT</sub>. Overall, both the tRNA structure and the tRNA:aIF2 $\gamma$  interface are constrained in IC1-P<sub>IN</sub> (Fig. 4a). Remarkably, these structural constraints resemble those observed during accommodation of the EF1A:tRNA complex in the A site upon codon recognition<sup>33,34</sup>. The remodeling of the TC is accompanied by the loosening of some contacts between aIF2 and the 30S subunit such as those involving the long L2 loop of aIF2 $\gamma$ -DII and the aIF2 $\alpha$ -D3 domain (Fig. 2b). Moreover, concomitant with its adjustment within the P site, the contacts of the tRNA with aIF2 $\alpha$ -D3 are lost. Accordingly, the electron density of aIF2 $\alpha$ -D3 appears weakened in IC1-P<sub>IN</sub>. In the same view, the two mobile wings of aIF2 (aIF2 $\alpha$ -D12 and core of aIF2 $\beta$ ) have moved and their fits in the electron density become poorly accurate (Fig. 2b and Supplementary Fig. 4b). The inside of the L-shaped tRNA, occupied by aIF2 $\beta$  in IC0-P<sub>OUT</sub> is now occupied by aIF1 (Fig. 2). The basic loop of aIF1 moves to allow tRNA accommodation and is now inserted between h44 and the anticodon loop (Supplementary Fig. 4c). The density corresponding to aIF1 appears less strong than in IC0-P<sub>OUT</sub> suggesting reduced interactions in IC1-P<sub>IN</sub> after start codon recognition. Accordingly, contacts between aIF1 and aIF2 seem less tight than in IC0-P<sub>OUT</sub> (Fig. 2). The two IC1-P<sub>IN</sub> and IC0-P<sub>OUT</sub> structures also differ in the orientation of the 30S head with respect to the body. The observed head movement is similar to that described for py48S when compared with 40S:eIF1:eIF1A<sup>22</sup>, and corresponds to a clockwise rotation around h28 of about 6° from IC0-P<sub>OUT</sub> to

IC1-P<sub>IN</sub> (Fig. 5a and Supplementary Fig. 5). By analogy with py48S, the head rotation would allow interaction of G<sub>1298</sub> and A<sub>1299</sub> of h29 with the minor groove G:C base pairs whereas h31 re-positioning would avoid clashes with the anticodon of the tRNA.

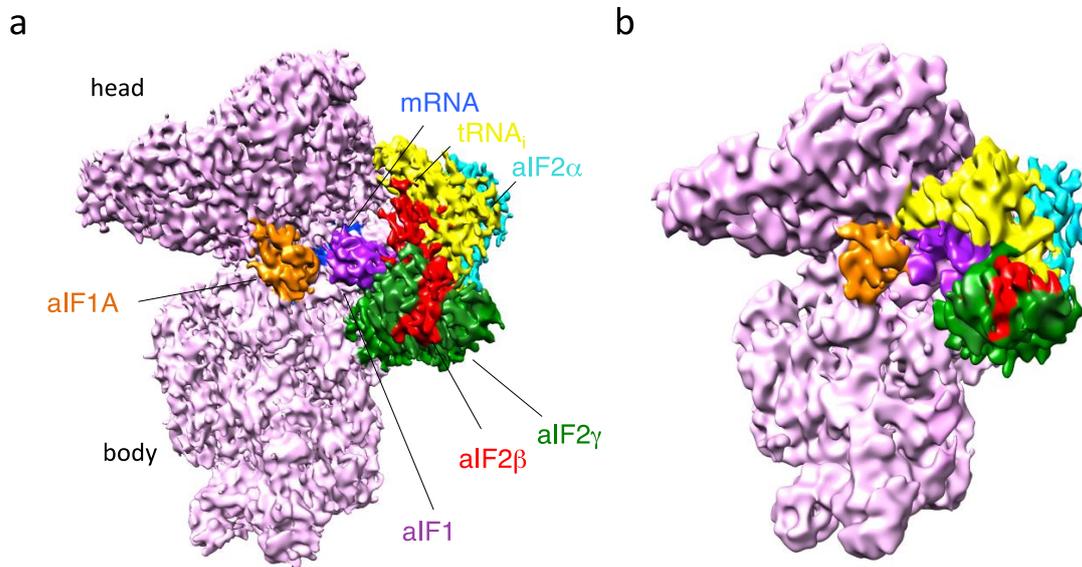
## DISCUSSION

Comparison of IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub> highlights the structural transition that occurs upon start codon recognition. The tRNA accommodation can be described as an isomerization of the IC0-P<sub>OUT</sub> complex whereby the tRNA molecule, surrounded by the  $\alpha$  and  $\beta$  subunits of aIF2, swings into the P site using the 3' methionyl-adenosine as a pivot point (Fig. 5 and Supplementary movie). This movement is accompanied by the head rotation allowing locking of the tRNA anticodon stem in the P site. Because the 3' methionyl-adenosine remains bound to aIF2 $\gamma$ , the establishment of the codon:anticodon interaction induces a structural constraint in the TC. This suggests that aIF2 exerts a spring force on the tRNA adjusted in the P site, which maintains the tRNA in the “out” conformation if there is no correct codon in the mRNA available. Upon recognition of the AUG start codon and formation of correct codon-anticodon pairing, the spring force on the tRNA is compensated and the tRNA remains positioned in the P site (Fig. 5). This represents a novel view of the contribution of aIF2 in the specificity of start codon selection. It appears that, during isomerization, the weak interaction of the  $\beta$  subunit with the tRNA is replaced by interaction with aIF1. This may give sense to the structural homology between aIF1 and aIF2 $\beta$ <sup>28,35</sup>. Finally, the steric constraint imposed by the tRNA pairing at the P site destabilizes aIF1, as shown by its weaker electron density in IC1-P<sub>IN</sub> as compared to IC0-P<sub>OUT</sub> (Fig. 2). By releasing the contacts of aIF1 with the switch regions of aIF2 $\gamma$ , this destabilization would in turn contribute to Pi release and to dissociation of aIF2-GDP.

The occurrence of the spring force mechanism in archaea raises the question of whether it could be a general mechanism in eukaryotes during the start codon selection process. Interestingly,

the rather slow rate of mRNA scanning during CAP-dependent initiation (5-10 nt/s<sup>36</sup>) makes it likely that a constrained state of the TC is also pivotal in eukaryotes. Because constrained states are intrinsically less stable, it is possible that the previously described py48S closed complex<sup>14</sup> represents a later step of the process in which eIF2 has been released from its constrained position on h44 (Fig. 4b). Indeed, in py48S-closed, the TC is unconstrained and no contact between aIF2 $\gamma$  and the ribosome are visible. Here, the use of a hyperthermophilic model organism may have helped trapping the constrained complex as an important intermediate. The contacts between aIF2 and helix h44, as well as the contacts between aIF2 and aIF1, although not observed in previously determined structures<sup>13,14,22</sup>, would make sense in eukaryotic complexes. As a matter of fact, labeling experiments have suggested close vicinity of eIF2 $\gamma$ -DIII and h44 in yeast<sup>18</sup>. Moreover, biochemical studies have proposed that eIF1 release triggers the release of the Pi group from eIF2<sup>7</sup>. In this view, it is remarkable that the regions of contact between aIF2 $\gamma$  and aIF1, as identified here, are both the targets of Sui<sup>-</sup> mutations in *S. cerevisiae*<sup>8,37,38</sup> suggesting that such interactions may be at the origin of communication between eIF2 and eIF1 during eukaryotic translation initiation. Taken together, the novel constrained P<sub>IN</sub> state represents a new key intermediate that explains a mechanism of mRNA scanning based on a spring force concept involving aIF2-assisted codon/anticodon probing by the initiator tRNA. This provides a novel view of the central role that e/aIF2 plays in start codon selection.

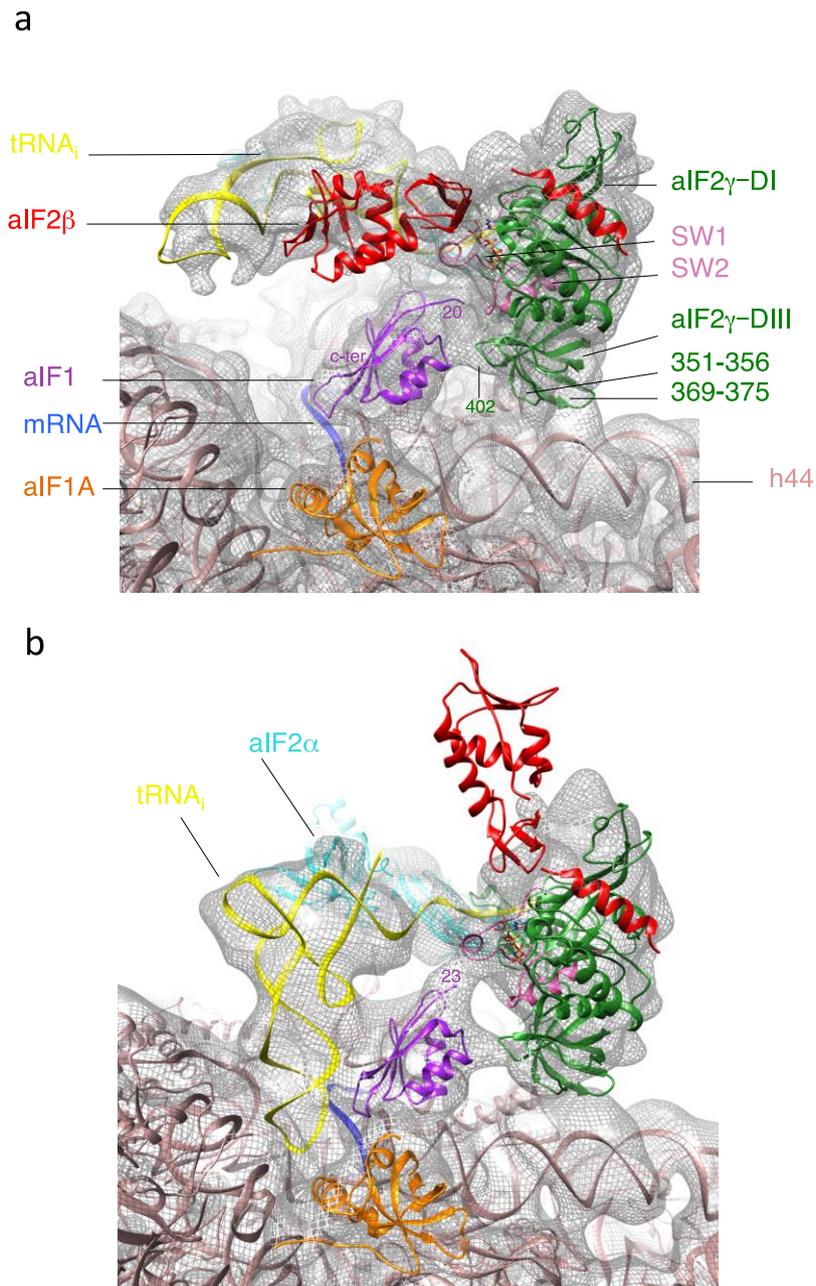
## Figures



**Figure 1: Cryo-EM maps of archaeal initiation complexes.**

(a) Cryo-EM map of the IC0-P<sub>OUT</sub> complex at 5.3 Å resolution. Regions of the map are colored by component to show the 30S subunit (pink), aIF1A (orange), aIF1 (magenta), Met-tRNA<sub>i</sub><sup>Met</sup> (yellow), mRNA (blue), aIF2γ (green), aIF2β (red) and aIF2α (cyan). The same color code is used in all figures.

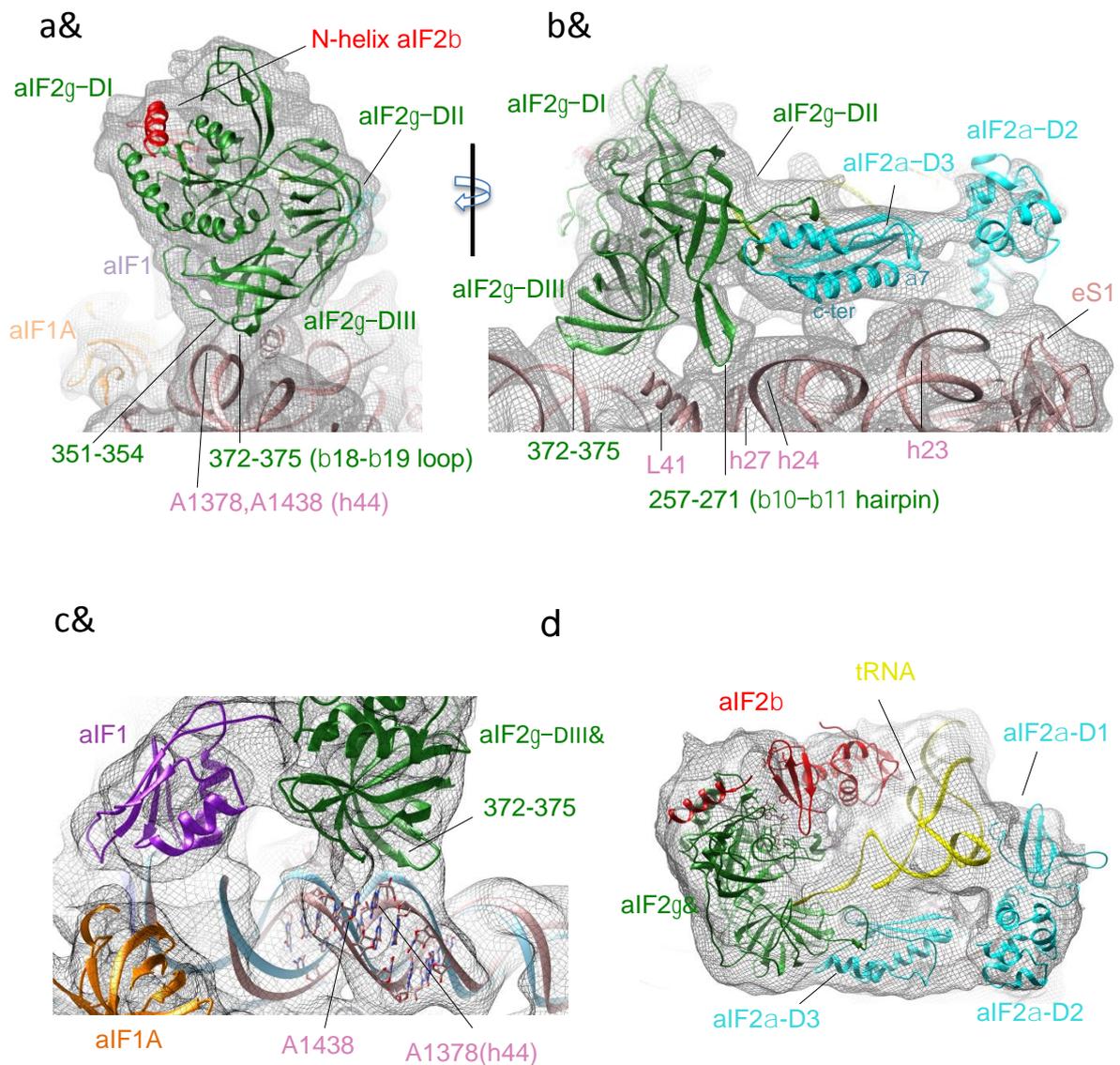
(b) Cryo-EM map of the IC1-P<sub>IN</sub> complex at 7.5 Å resolution.



**Figure 2: Views of the P site region in the cryo-EM structures of IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub>.**

(a) IC0-P<sub>OUT</sub>. The cryo-EM map has been low-pass filtered to 8.5 Å resolution. The models used to account for the electron density correspond to the structures of *Pyrococcus furiosus* 30S (PDB entry 4V6U<sup>39</sup>), of *Sulfolobus solfataricus* TC (PDB entry 3V11<sup>24</sup>), of Pa-aIF1A (PDB entry 4MNO) and of *Methanocaldococcus jannaschii* aIF1 (PDB entry 4MO0) (Supplementary Fig. 4). Note that the residue numbering used in the text refers to that of the fitted components. Switch 1 and 2 regions in aIF2γ are colored in light pink. GDPNP is shown in sticks. This view also highlights contacts between aIF1 and aIF2γ. One contact involves the β20-β21 loop (residues 402-404) of aIF2γ with the aIF1 α2 helix. A second area of density connects the two switch regions of aIF2γ and the N-terminal domain of aIF1.

(b) IC1-P<sub>IN</sub>. The density map has been low-passed filtered to 12 Å resolution. Note that positioning of the core domain of aIF2β is only tentative.



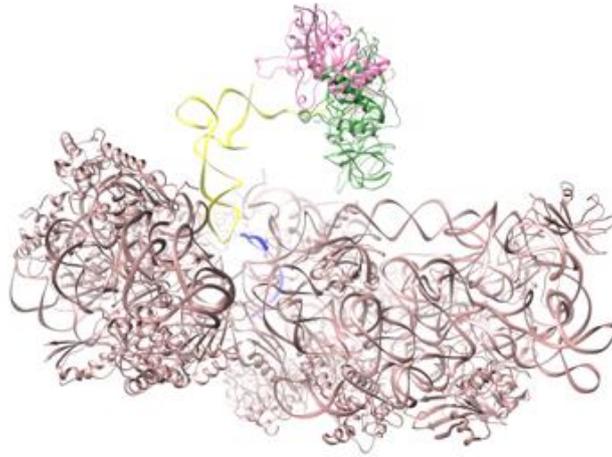
**Figure 3: Contacts between aIF2 and the 30S ribosomal subunit as observed in IC0-P<sub>OUT</sub>.**

(a) aIF2γ–DIII is in contact with the h44 bulge in 16S RNA. These contacts involve two regions, 351-354 and 372-375, as labeled in green.

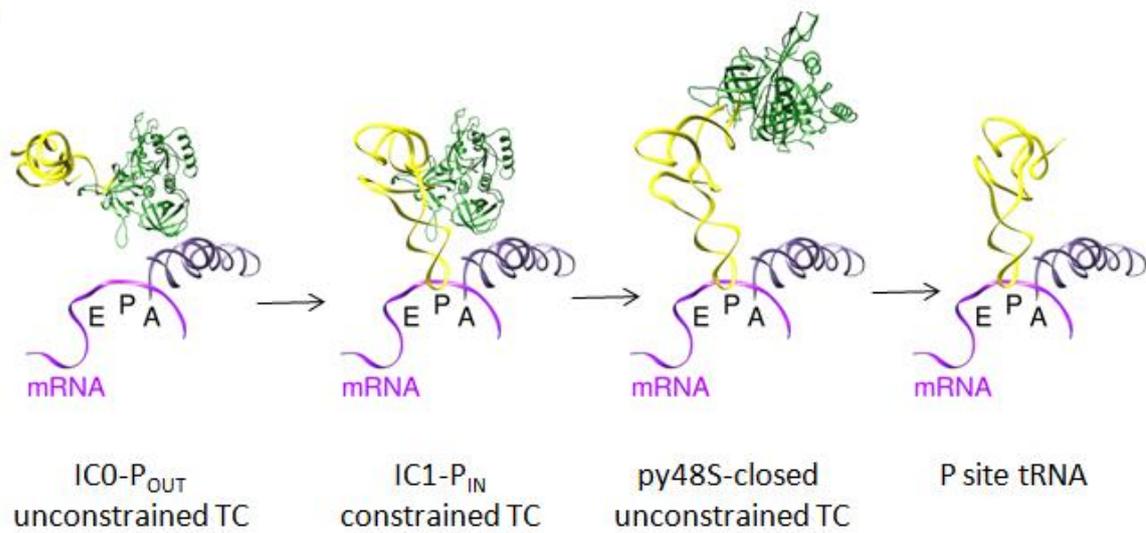
(b) The view is rotated by c.a. 90° as compared to (a) in order to highlight the contact of the L2 loop of aIF2γ–DII with h24 and h27, at the hinge between the decoding center and the platform. Contacts between aIF2α and the ribosome are also visible (see text for details).

(c) Fit of h44 (pink) into IC0-P<sub>OUT</sub> map is shown. The view shows that the structure of h44 from Pf-30S subunit (blue, PDB entry 4V6U<sup>39</sup>) had to be remodeled to account for the density. (d) View showing how aIF2 subunits wrap the initiator tRNA in IC0-P<sub>OUT</sub>. The cryo-EM map is low-pass filtered to 8.5 Å.

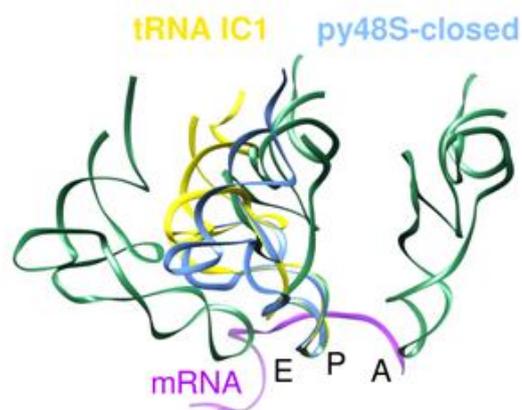
a



b



c



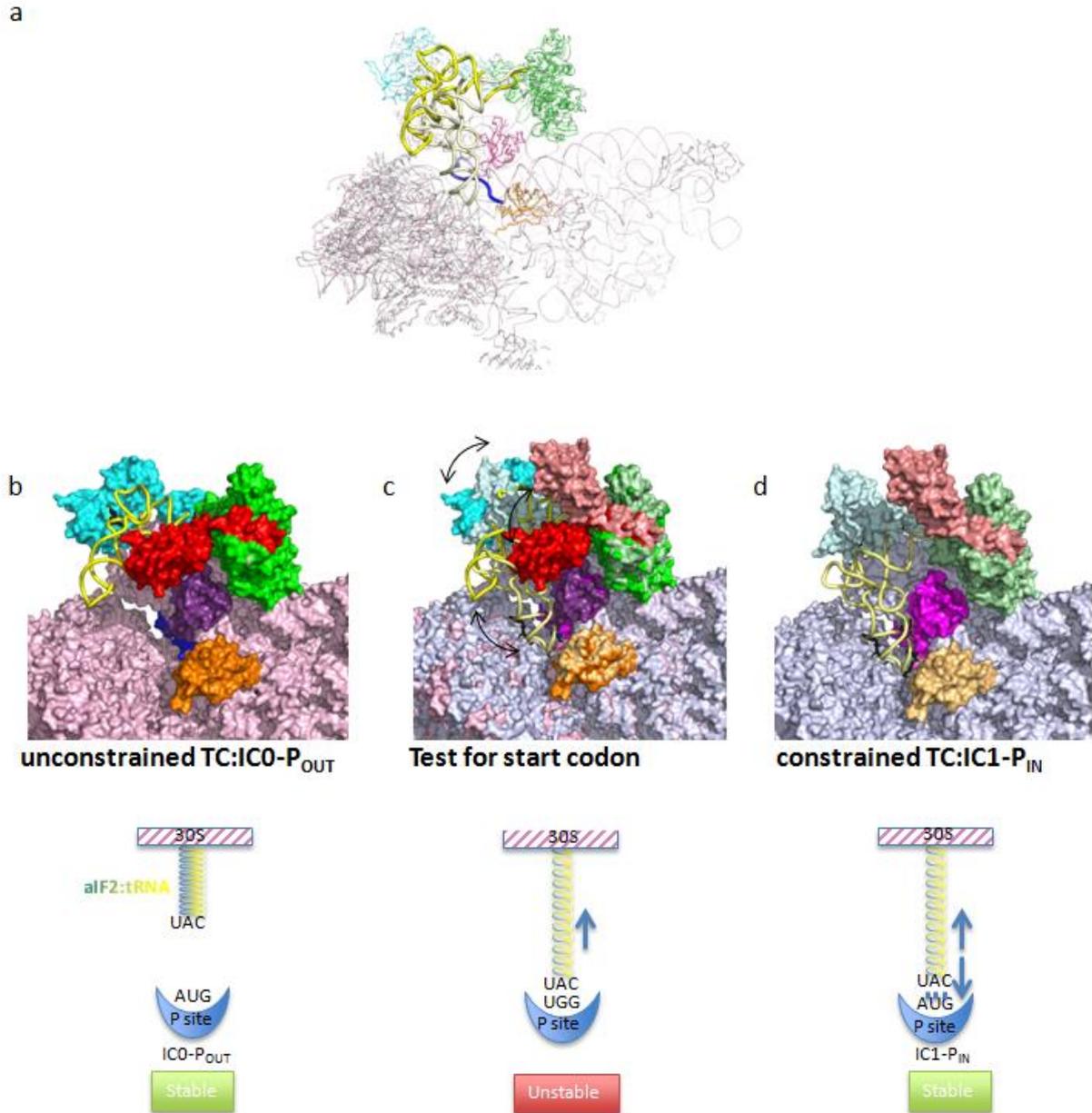
**Figure 4: Adjustment of the initiator tRNA in the P site.**

(a) The tRNA in the free TC (PDB entry 3V11) has been superimposed on the tRNA in IC1-P<sub>IN</sub>. The observed in IC1-P<sub>IN</sub> (green). IC1-P<sub>IN</sub> is colored as follows; 30S subunit in light pink, tRNA in yellow,

mRNA in blue and aIF2 $\gamma$  in green. For the sake of clarity other components are not represented. The view illustrates that the TC is constrained in IC1-P<sub>IN</sub>.

(b) Cartoon comparing the position of aIF2 $\gamma$ -tRNA in different ribosomal initiation complexes. h44 (purple) and P site tRNA (yellow) from the 70S *T. thermophilus* complex (PDB entry 4V51<sup>32</sup>) are represented on the right for reference. mRNA (magenta) is from PDB entry 1JG0<sup>40</sup>. The relative positions of aIF2 $\gamma$ -tRNA in IC0-P<sub>OUT</sub>, in IC1-P<sub>IN</sub> and in py48S-closed were deduced from superimposition of the body of the small ribosomal subunit of each complex on that of the *T. thermophilus* 70S. aIF2 $\gamma$  is colored in green. The view illustrates that the TC is constrained in IC1-P<sub>IN</sub> whereas it is unconstrained in IC0-P<sub>OUT</sub> and py48S-closed complex.

b-Comparison of IC1-P<sub>IN</sub> tRNA with P site tRNAs in 70S (PDB entry 4V51<sup>32</sup>) or in py48S-closed PIC (PDB entry 3JAP<sup>14</sup>). A and E site tRNAs are from PDB entry 4V51, mRNA is from PDB entry 1JG0<sup>40</sup>.



**Figure 5: Model of the role of aIF2 in start codon selection.**

(a) The positions of the TC in IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub> are compared after superimposition of the 30S bodies. Color code for IC1-P<sub>IN</sub> is as follows: ribosome in gray, tRNA in light yellow, aIF2 $\gamma$  in light green cartoons, aIF2 $\alpha$  in light cyan, aIF1 in light pink, aIF1A in light orange and mRNA in blue. Color code for IC0-P<sub>OUT</sub> is as follows: 30S head in pink, tRNA in yellow, aIF2 $\gamma$  in green, aIF2 $\alpha$  in cyan, aIF1 in magenta, aIF1A in orange and mRNA in blue. For the sake of clarity, aIF2 $\beta$  is not represented. The view shows that aIF2 $\gamma$  does not significantly move during the IC0-P<sub>OUT</sub> to IC1-P<sub>IN</sub> transition. Therefore, the TC is distorted in IC1-P<sub>IN</sub>. The view also shows that accommodation of the tRNA in the P site is accompanied by a rotation of the 30S head.

(b) Surface representation of IC0-P<sub>OUT</sub>: color code is same as in panel (a). The TC has the same conformation as unbound to 30S and is therefore unconstrained.

(c) Surface representation of IC0-P<sub>OUT</sub> superimposed on IC1-P<sub>IN</sub>. Color code is same as in panel (a). The tRNA can explore the P site although this results in a constrained TC conformation.

(d) Surface representation of IC1-P<sub>IN</sub>. If an AUG codon is found in the P site, the constrained conformation of the TC is stabilized by base pairing.

The lower part of the figure shows schemes illustrating the spring force model.

## ACCESSION CODES

The EMDataBank accession numbers for the EM maps reported in this paper are EMD-8148 (IC0-P<sub>OUT</sub>) and EMD-8149 (IC1-P<sub>IN</sub>). Coordinates of the models fitted in the maps have been deposited in the Protein Data Bank (accession numbers 5JB3 for IC0-P<sub>OUT</sub>, 5JBH for IC1-P<sub>IN</sub>).

## ACKNOWLEDGMENTS

We thank F. Weis, R. Gillet, S. Jonic, A. Myasnikov for valuable discussions on electron microscopy and J.F. Menetret for help with data collection at the preliminary stages of this study. We also gratefully acknowledge R. *Matadeen* and S. de Carlo (FEI Company) for image acquisition at the National Center for Electron Nanoscopy in Leiden (*NeCEM*) and helpful advice. This work was partly supported by the Physique et Chimie du Vivant program of Agence Nationale de la Recherche (MASTIC project) and by the French Infrastructure For Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01. AM was supported by a “Bourse Monge” doctoral fellowship from Ecole polytechnique.

## AUTHOR CONTRIBUTIONS

PDC, ES, YM designed the research; PDC, YM, ES, CLS, AM, EL, LC, carried out the experiments; all authors analyzed the data; PDC, ES, YM, BPK wrote the paper.

## METHODS

### Growth of *P. abyssi* and preparation of ribosomal particles

*P. abyssi* cells were grown under anaerobic conditions<sup>41</sup>. Mass cultivations were carried out at 90°C in a 160 L fermenter (Laboratoire de microbiologie industrielle UMR FARE 614, Reims, FRANCE). The cells were harvested by centrifugation, aliquots were flash-frozen in liquid nitrogen and stored at -80°C until further use.

To isolate archaeal ribosomes, 3 g of cell pellets were dissolved in 16 mL of buffer A (10mM Hepes pH 7.5, 100mM NH<sub>4</sub>Cl, 10.5 mM MgAcetate, 0.1 mM EDTA, 6 mM 2-mercaptoethanol, 0.1 mM PMSF, 0.1 mM Benzamidine) at 4°C, mixed with 16 mL of glass beads (5 µm, SIGMA), and disrupted by using a Vibrogen-Zellmühle apparatus (Edmund Bühler, Fisher Scientific). After extensive washing of the glass beads and centrifugations, the crude extract (34 mL) was loaded onto a 35 mL sucrose cushion (10 mM Hepes pH 7.5, 1.1 M sucrose, 1 M NH<sub>4</sub>Cl, 10.5 mM MgAcetate, 0.1 mM EDTA, 6 mM 2-mercaptoethanol, 0.1 mM PMSF, 0.1 mM benzamidine). After centrifugation at 235,000 g for 19h at 4°C, the pellet was dissolved in 3 mL of buffer A and loaded onto a Sephacryl S400 column (8 mm x 60 cm, GE Healthcare). Fractions containing ribosomal particles were pooled, dialyzed for 3 h against buffer B (same as buffer A but 2.5 mM MgAcetate) and then loaded on the top of sucrose gradients (10-30 % sucrose in buffer B made using BioComp Gradient Master, Fredericton, Canada). After a 19 h centrifugation at 70,000g (SW32.1 rotor, Beckman), the gradients were fractionated. Fractions containing 30S subunits were pooled and the magnesium concentration was increased to 10.5 mM. After centrifugation (220,000g, 20 h), the pellet was briefly rinsed with buffer A to remove residual sucrose and then dissolved in buffer A. About 400 µL of 30S subunits at 2.7 µM were usually obtained. The 30S subunits obtained using this procedure were quite pure, as shown by native and SDS-PAGE analyses (Supplementary Fig. 1). However, a contaminating band was systematically visible on SDS gels. This band was identified as phosphoenol pyruvate synthase by

peptide mass fingerprinting. This high molecular weight protein results from the octameric assembly of a 91 kDa monomer.

### **Purification of initiation factors and of aminoacyl-tRNAs**

#### *Pa-aIF2*

Cultures (250 mL) of cells overproducing each of the three subunits (native aIF2 $\gamma$ , N-terminally tagged versions of aIF2 $\alpha$  and aIF2 $\beta$ ) were harvested, mixed in 40 mL of buffer C (500 mM NaCl, 10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM PMSF, 0.1 mM benzamidine) and disrupted by sonication. After centrifugation, the supernatant was heated for 10 min at 80°C. After pelleting the nonthermostable proteins by centrifugation, the supernatant was directly loaded onto a Q-Hiload column (10 mm x 4 cm; GE Healthcare) equilibrated in buffer C. The flow through was recovered and diluted two fold with buffer D (10 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1 mM PMSF, 0.1 mM benzamidine). The sample was then loaded onto an S-Hiload column (10 mm x 4 cm; GE Healthcare) equilibrated in buffer D containing 250 mM NaCl. Assembled heterotrimer was eluted by applying a step of buffer D containing 700 mM NaCl. The recovered sample was then concentrated to 1 mL. Final heterotrimer preparation was obtained after purification by molecular sieving on a Superdex 200 HR column (10 mm x 30 cm, GE Healthcare). At the end of purification, the protein was concentrated and 1 mM Gpp(NH)p-Mg<sup>2+</sup> was added. Flash-frozen samples were stored at -80°C.

#### *Pa-aIF1*

The gene encoding aIF1 from *P. abyssi* was amplified from genomic DNA and cloned into pET15b1pa to produce an N-terminally his-tagged version of the factor. A culture (250 mL) of cells overproducing the tagged version of aIF1 was harvested, re-suspended in 10 mL of buffer D plus 100 mM NaCl and heated for 10 min at 70°C. After pelleting non-thermostable proteins by centrifugation, the supernatant was directly loaded onto a Talon affinity column (10 mm x 4 cm; Clontech) equilibrated in buffer D plus 100 mM NaCl. The tagged protein was eluted with a step of buffer D containing 125 mM imidazole. After dialysis against buffer D containing 250 mM NaCl, the preparation was passed through a Q-Hiload column equilibrated in buffer A containing 100 mM NaCl. This step was repeated twice in order to efficiently remove nucleic acids. After concentration, the protein was flash-frozen and stored at -80°C. Routinely, 4 mg of purified protein were obtained from a 250 mL culture.

#### *Pa-aIF1A*

The gene encoding aIF1A from *P. abyssi* was amplified from genomic DNA and cloned into pET3alpa to produce a native version of the factor. A culture (250 mL) of cells overproducing the native version of aIF1A was harvested, resuspended in 10 mL buffer A containing 500 mM NaCl and disrupted by sonication. After centrifugation, the supernatant was heated for 10 min at 80°C. After pelleting the nonthermostable proteins by centrifugation, the supernatant was directly loaded onto a Q-Hiload column (10 mm x 4 cm; GE Healthcare) equilibrated in buffer A plus 500 mM NaCl. The flow through was recovered, diluted twice to obtain a final concentration of 250 mM NaCl and loaded onto an S-Hiload column (10mm x 4 cm; GE Healthcare) equilibrated in buffer A containing 250 mM NaCl. aIF1A was eluted by applying an NaCl gradient (250 to 800 mM) in buffer A. The recovered sample was then concentrated, flash-frozen and stored at -80°C. Routinely, 4 mg purified protein were obtained from a 250 mL culture.

### *Preparations of methionylated tRNAs.*

tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub> was produced in *E. coli* from a cloned gene, purified and aminoacylated as described<sup>42</sup>. As previously shown, Met-tRNA<sub>f</sub><sup>Met</sup> from *E. coli* is 79% identical to Pa-tRNA<sub>i</sub><sup>Met</sup> and the two tRNAs bind Pa-aIF2 with almost the same affinity<sup>25</sup>. Moreover, SAXS data have shown that an aIF2-tRNA complex formed with either Met-tRNA<sub>f</sub><sup>Met</sup> from *E. coli* or archaeal Met-tRNA<sub>i</sub><sup>Met</sup> display identical diffusion curves, which strongly suggest similar structures<sup>29</sup>.

### *Purification of IC*

The initiation complex was assembled as follows: 864 picomoles of Pa-30S were mixed to a 2-fold excess of Pa-aIF2 (1728 picomoles), 5-fold excess of Pa-aIF1 (4322 picomoles), Pa-aIF1A, met-tRNA<sub>f</sub><sup>Met</sup> A<sub>1</sub>-U<sub>72</sub>, aEF1A mRNA (A<sub>(-17)</sub>UUUGGAGGUGAUUUAAA<sub>(+1)</sub>UGCCAAAG<sub>(+9)</sub>, ThermoScientific) in the presence of GDPNP to a final concentration of 1mM. The mixture was incubated 5 minutes at 51°C in IC buffer (10mM Hepes pH 7.5, 100mM NH<sub>4</sub>Cl, 10mM MgAcetate, 3mM 2-mercaptoethanol) and then 1.75 h at room temperature. The IC complex was purified using Talon affinity column and standard procedures. The recovered complex (Supplementary Fig. 1) was dialyzed against IC buffer and concentrated to 2.6 μM using centricon 30, and then stored at -80°C after flash freezing in liquid nitrogen. The presence of all components was checked using SDS-PAGE (Supplementary Fig. 1).

### **Electron microscopy**

Before spotting onto the grids, an excess of mRNA, aIF1A, TC (each 1:1 regarding IC) and GDPNP (0.5mM) was added and the IC was heated 1 min at 50°C. After heating, the samples were diluted to a final concentration of c.a. 0.2 μM in IC buffer containing 0.5 mM GDPNP and immediately used for grid preparations. Carbon-coated Quantifoil R2/2 grids (Quantifoil Micro Tools GmbH) were

blotted with 4  $\mu\text{L}$  of the purified samples and vitrified in liquid ethane at 100 K on a Vitrobot Mark IV (FEI). Low-dose images were collected at a calibrated magnification of 129629x on a FEI Titan Krios with a Cs corrector operating at 300 kV, using a 4k x 4k FEI Falcon II camera (Netherlands Centre for Electron Nanoscopy, Leiden). The pixel size was 1.12  $\text{\AA}$  on the camera. Defocus values ranged from -0.5  $\mu\text{m}$  to -2.8  $\mu\text{m}$ . All images were recorded using the EPU software (FEI) for automatic data collection.

### Image processing

Low-dose images were taken with a total dose of  $44\text{e}^-/\text{\AA}^2$  summed on 16 frames. The first frame and the last 8 frames were discarded to minimize beam induced movement and irradiation. The remaining frames were first aligned with DRIFTCORR<sup>43</sup>, averaged and then sorted according to best power spectrum and ice quality. Particles from the selected images ( $\sim 2100$ ) were automatically picked with the RELION package<sup>21</sup>. An initial dataset of  $\sim 192,000$  particles was sorted using several rounds of classical 2D-classification protocols implemented in RELION. Then, a 3D-classification in RELION using a 70  $\text{\AA}$  filtered map of the archaeal 30S (extracted from the *Pyrococcus furiosus* 70S structure<sup>39</sup>) helped to isolate several well-resolved classes that were used for further 3D-classification and 3D-refinement (Supplementary Fig. 1d). Two final datasets of  $\sim 12600$  and  $\sim 5500$  particles were used to get the final 3D reconstructions called IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub>, respectively. The resolutions of the final three-dimensional reconstructions were estimated, using gold-standard Fourier Shell Correlation, to 9.5  $\text{\AA}$  and 12.6  $\text{\AA}$  for IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub>, respectively. A movie processing approach<sup>44</sup> was also tempted in RELION but did neither increase the maps quality nor their resolution. Finally, a post-processing approach also implemented in RELION gave higher resolution maps of 5.34  $\text{\AA}$  and 7.5  $\text{\AA}$  (Supplementary Fig. 1e) for IC0-P<sub>OUT</sub> and IC1-P<sub>IN</sub>, respectively. Reported overall resolutions were calculated with the gold-standard FSC = 0.143 criterion<sup>21</sup> and were corrected for the effects of a soft mask on the FSC curve with high-resolution noise substitution<sup>45</sup>.

The final density maps were corrected for the modulation transfer function of the detector and sharpened by applying a negative B factor that was estimated using automated procedures<sup>46</sup>. Sharpening B-values were  $-128 \text{ \AA}^2$  for IC0-P<sub>OUT</sub> and  $-208 \text{ \AA}^2$  for IC1-P<sub>IN</sub>.

To assess the quality of the final reconstructions, a local resolution map for each reconstruction was also computed: information up to  $4 \text{ \AA}$ , for the core of the 30S subunit and between  $6 \text{ \AA}$  and  $15 \text{ \AA}$  for the translation initiation factors binding regions is available in both complexes (Supplementary Fig. 1g,h). Such decrease in resolution as distance with the core 30S increases was already observed in other initiation complexes<sup>22</sup>. Because of this, for better clarity in the figures showing factors, maps were generally low-pass filtered to  $8.5 \text{ \AA}$  (IC0-P<sub>OUT</sub>) or  $12 \text{ \AA}$  (IC1-P<sub>IN</sub>).

### Model building and refinement

#### IC0-P<sub>OUT</sub>:

The structures of the head and body parts of the 30S subunits were extracted from the *P. furiosus* 70S coordinates<sup>39</sup> (PDB entry 4V6U). Head and body were manually placed independently into the electron density map using Chimera, before automatically refining the position in the same program. Messenger RNA was clearly visible in the density, including interaction of the SD sequence with the anti-SD (Supplementary Fig. 3a). mRNA was modeled in Coot<sup>47</sup> using the structure of the *T. thermophilus* 70S ribosome complexed with mRNA and tRNAs (PDB entry 1JGO<sup>40</sup>) as a guide. The start codon was positioned in the P site. mRNA was not visible from base +5 to base +9 (as in yeast 48S complex, PDB entry 3JAQ) at the 3' extremity. Bases -17 and -16 at the 5' end of mRNA were not visible. Isolated PDB models of initiation factors and tRNA were manually placed in the electron density. The crystallographic model of the ternary initiation complex aIF2:GDPNP:tRNA<sup>24</sup> (PDB entry 3V11) could be readily fitted in the electron density with few modifications (see text and Figs. 2a and

3d). Finally, a few manual model corrections were performed in Coot<sup>47</sup> to better fit the electron density maps. These corrections concerned the following regions: 16S rRNA connection between head and body, h44 bulge, a1F1 basic loop, h44 near a1F1A (as described below). The obtained model was then refined in Refmac<sup>48</sup> using data up to 5.8 Å resolution. Tight geometric and external restraints (generated using ProSmart and LIBL) were maintained throughout refinement in order to minimize overfitting, as described<sup>49-51</sup>. According to the relatively modest resolution, an overall B value was used for the whole model. Average Fourier shell correlation (FSC) was monitored in Refmac. For cross-validation against overfitting, the atoms of the final model were randomly displaced (with a maximum deviation of 0.5 Å), and a refinement procedure was performed against the maps that were reconstructed from only one of the two independent halves of the data used in our gold-standard FSC procedure. FSC curves were calculated between the resulting model and the half map against which it had been refined (FSCwork) as well as the FSC curve between the model and the other half map (FSCtest)<sup>49</sup> (Supplementary Fig.1f). The average FSC for the final model against the 5.8 Å resolution data used in refinement was 0.56 (Rfactor 0.42). Correlation coefficient between the map simulated from the model (5.34 Å resolution) and the experimental map was 82.1 %, as determined using Chimera<sup>52</sup>. Further attempts to improve this value by loosening restraints during refinement systematically led to overfitting. Therefore, at the present resolution, Refmac was essentially used to fix the small clashes and to improve geometry of the initial 30S unrefined cryo-EM model. The number of bad bonds dropped from 5.93 % to 0.06 %, and the number of bad angles from 6.93 % to 0.58 %.

#### IC1-P<sub>IN</sub>:

The refined head, body and mRNA from IC0-P<sub>OUT</sub> were fitted within IC1 map using Chimera. Positions of a1F1 and a1F1A were manually adjusted in the electron density. According to its known structural organization<sup>28,53</sup>, a1F2 was placed as three rigid bodies: (1) the  $\gamma$  subunit bound to domain 3 of the  $\alpha$  subunit and to the N-terminal helix of the  $\beta$  subunit, (2) domains 1 and 2 of the  $\alpha$  subunit

and (3) the core domain of the  $\beta$  subunit. The model of the tRNA was built using the structure of the P site initiator tRNA as observed in 4V51<sup>32</sup>. The tRNA was cut into two parts in the hinge region (G<sub>26</sub>A<sub>44</sub>). The two parts were fitted independently. Start codon and anticodon tRNA bases were adjusted to allow base pairing. Moreover, the acceptor CCA end was remodeled to allow fitting of the methionylated 3'adenosine in its pocket on the aIF2 $\gamma$ -DII domain. Finally, the connection between the two tRNA halves was manually adjusted. As a final step, the basic loop of aIF1 was rebuilt to better fit the density. Because of the modest resolution of IC1-P<sub>IN</sub>, Refmac refinement systematically led to overfitting and was therefore not applied to the final model. Correlation coefficient between the map simulated from the model (7.5 Å resolution) and the experimental map was 88.0 %, as determined using Chimera<sup>52</sup>.

## REFERENCES

1. Hershey, J.W., Sonenberg, N. & Mathews, M.B. Principles of translational control: an overview. *Cold Spring Harb. Perspect. Biol.* **4**, a011528 (2012).
2. Hinnebusch, A.G. & Lorsch, J.R. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb. Perspect. Biol.* **4**, a011544 (2012).
3. Jackson, R.J., Hellen, C.U. & Pestova, T.V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell. Biol.* **11**, 113-127 (2010).
4. Asano, K. et al. Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO J.* **20**, 2326-2337 (2001).
5. Majumdar, R., Bandyopadhyay, A. & Maitra, U. Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *J. Biol. Chem.* **278**, 6580-6587 (2003).
6. Pestova, T.V., Borukhov, S.I. & Hellen, C.U. Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* **394**, 854-859. (1998).
7. Algire, M.A., Maag, D. & Lorsch, J.R. Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol. Cell* **20**, 251-262 (2005).

8. Cheung, Y.N. et al. Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo. *Genes Dev.* **21**, 1217-1230 (2007).
9. Huang, H.K., Yoon, H., Hannig, E.M. & Donahue, T.F. GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. *Genes Dev.* **11**, 2396-2413 (1997).
10. Unbehaun, A., Borukhov, S.I., Hellen, C.U. & Pestova, T.V. 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. (2004).
11. Aitken, C.E. & Lorsch, J.R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **19**, 568-576 (2012).
12. Voigts-Hoffmann, F., Klinge, S. & Ban, N. Structural insights into eukaryotic ribosomes and the initiation of translation. *Curr. Opin. Struct. Biol.* **22**, 768-777 (2012).
13. Hashem, Y. et al. Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108-1119 (2013).
14. Llacer, J.L. et al. Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. *Mol. Cell* **59**, 399-412 (2015).
15. Lomakin, I.B. & Steitz, T.A. The initiation of mammalian protein synthesis and mRNA scanning mechanism. *Nature* **500**, 307-311 (2013).
16. Weisser, M., Voigts-Hoffmann, F., Rabl, J., Leibundgut, M. & Ban, N. The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat. Struct. Mol. Biol.* **20**, 1015-1017 (2013).
17. Saini, A.K., Nanda, J.S., Lorsch, J.R. & Hinnebusch, A.G. Regulatory elements in eIF1A control the fidelity of start codon selection by modulating tRNA<sup>i</sup>(Met) binding to the ribosome. *Genes Dev.* **24**, 97-110 (2009).
18. Shin, B.S. et al. Initiation factor eIF2 $\gamma$  promotes eIF2-GTP-Met-tRNA<sup>i</sup>(Met) ternary complex binding to the 40S ribosome. *Nat. Struct. Mol. Biol.* **18**, 1227-1234 (2011).
19. Hasenöhrl, D., Fabbretti, A., Londei, P., Gualerzi, C.O. & Bläsi, U. Translation initiation complex formation in the crenarchaeon *Sulfolobus solfataricus*. *RNA* **15**, 2288-2298 (2009).
20. Scheres, S.H. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406-418 (2012).
21. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519-530 (2012).
22. Hussain, T. et al. Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159**, 597-607 (2014).

23. Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A. & Ban, N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* **331**, 730-736 (2011).
24. Schmitt, E. et al. Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat. Struct. Mol. Biol.* **19**, 450-454 (2012).
25. Yatime, L., Schmitt, E., Blanquet, S. & Mechulam, Y. Functional molecular mapping of archaeal translation initiation factor 2. *J. Biol. Chem.* **279**, 15984-15993 (2004).
26. Stolboushkina, E. et al. Crystal structure of the archaeal translation initiation factor 2 in complex with a GTP analogue and Met-tRNA<sup>f</sup>(Met.). *J. Mol. Biol.* **425**, 989-998 (2013).
27. Yatime, L., Mechulam, Y., Blanquet, S. & Schmitt, E. Structural switch of the gamma subunit in an archaeal aIF2 alpha gamma heterodimer. *Structure* **14**, 119-128 (2006).
28. Yatime, L., Mechulam, Y., Blanquet, S. & Schmitt, E. Structure of an archaeal heterotrimeric initiation factor 2 reveals a nucleotide state between the GTP and the GDP states. *Proc Natl Acad Sci U S A* **104**, 18445-18450 (2007).
29. Naveau, M. et al. Roles of yeast eIF2alpha and eIF2beta subunits in the binding of the initiator methionyl-tRNA. *Nucleic Acids Res.* **41**, 1047-1057 (2013).
30. Pisarev, A.V. et al. Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev.* **20**, 624-636 (2006).
31. Fekete, C.A. et al. N- and C-terminal residues of eIF1A have opposing effects on the fidelity of start codon selection. *EMBO J.* **26**, 1602-1614 (2007).
32. Selmer, M. et al. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942 (2006).
33. Schmeing, T.M. et al. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688-694 (2009).
34. Valle, M. et al. Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J.* **21**, 3557-3567 (2002).
35. Conte, M.R. et al. Structure of the eukaryotic initiation factor (eIF) 5 reveals a fold common to several translation factors. *Biochemistry* **45**, 4550-4558 (2006).
36. Vassilenko, K.S., Alekhina, O.M., Dmitriev, S.E., Shatsky, I.N. & Spirin, A.S. Unidirectional constant rate motion of the ribosomal scanning particle during eukaryotic translation initiation. *Nucleic Acids Res.* **39**, 5555-5567 (2011).
37. Dorris, D.R., Erickson, F.L. & Hannig, E.M. Mutations in GCD11, the structural gene for eIF-2 gamma in yeast, alter translational regulation of GCN4 and the selection of the start site for protein synthesis. *EMBO J.* **14**, 2239-2249. (1995).

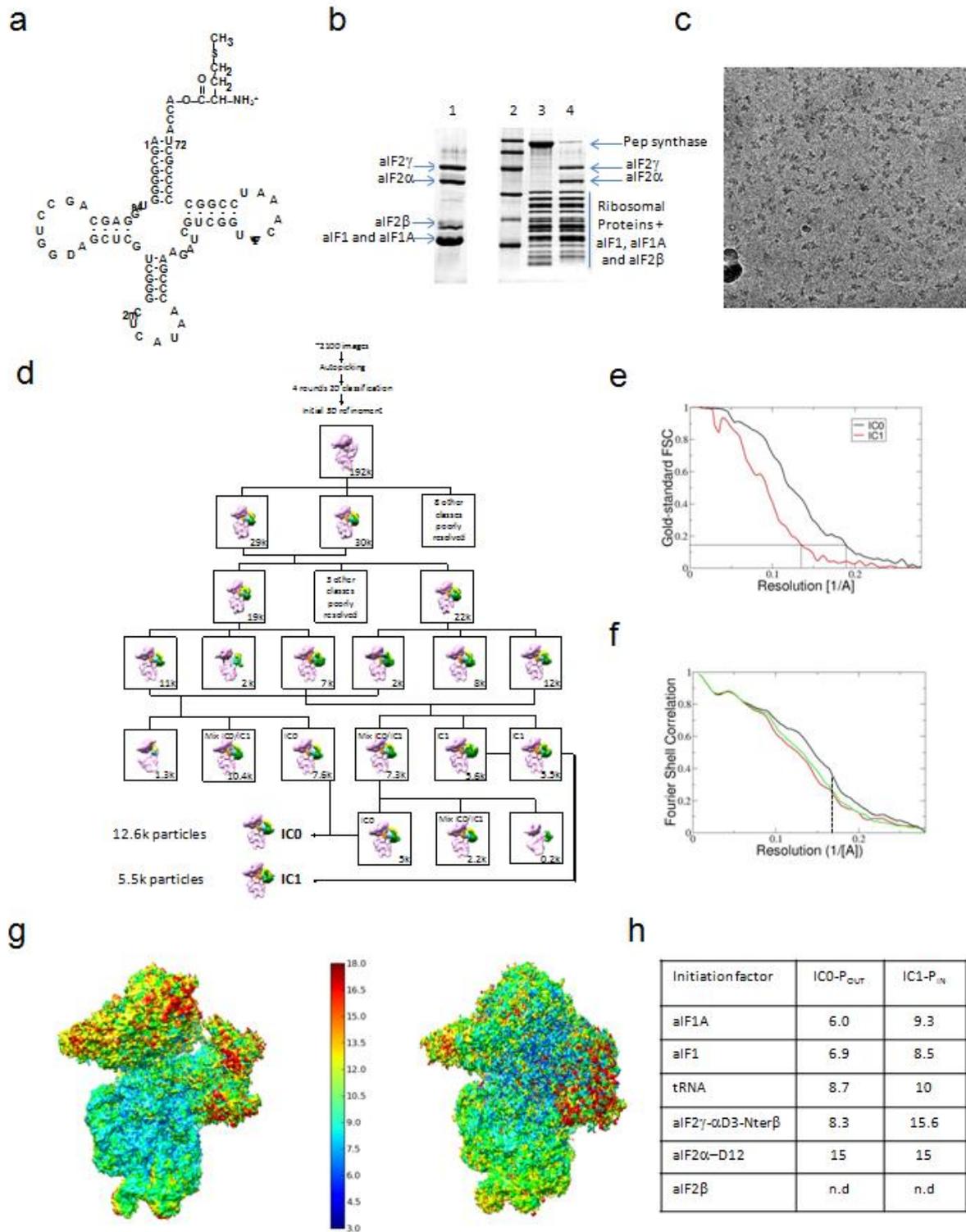
38. Martin-Marcos, P. et al. beta-hairpin loop of eIF1 mediates 40S ribosome binding to regulate initiator tRNA<sup>Met</sup> recruitment and accuracy of AUG selection in vivo. *J. Biol. Chem.*, **288**, 27546-27562 (2013).
39. Armache, J.P. et al. Promiscuous behaviour of archaeal ribosomal proteins: implications for eukaryotic ribosome evolution. *Nucleic Acids Res.* **41**, 1284-1293 (2013).
40. Yusupova, G.Z., Yusupov, M.M., Cate, J.H. & Noller, H.F. The path of messenger RNA through the ribosome. *Cell* **106**, 233-241 (2001).
41. Erauso, G. et al. *Pyrococcus abyssi* sp. nov. , a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* **160**, 338-349 (1993).
42. Mechulam, Y., Guillon, L., Yatime, L., Blanquet, S. & Schmitt, E. Protection-based assays to measure aminoacyl-tRNA binding to translation initiation factors. *Methods Enzymol.* **430**, 265-281 (2007).
43. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584-590 (2013).
44. Bai, X.C., Fernandez, I.S., McMullan, G. & Scheres, S.H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461 (2013).
45. Chen, S. et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24-35 (2013).
46. Rosenthal, P.B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **333**, 721-745 (2003).
47. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr.* **D66**, 486-501 (2010).
48. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255 (1997).
49. Amunts, A. et al. Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485-1489 (2014).
50. Brown, A. et al. Structure of the large ribosomal subunit from human mitochondria. *Science* **346**, 718-722 (2014).
51. Weis, F. et al. Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nat Struct Mol Biol* **22**, 914-919 (2015).
52. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612 (2004).

53. Stolboushkina, E. et al. Crystal structure of the intact archaeal translation initiation factor 2 demonstrates very high conformational flexibility in the alpha- and beta-subunits. *J. Mol. Biol.* **382**, 680-691 (2008).

## **Supplementary Information**

Initiator tRNA accommodation during translation initiation mRNA probing

Pierre-Damien Coureux, Christine Lazenec-Schurdevin, Auriane Monestier, Eric Larquet,  
Lionel Cladière, Bruno P Klaholz, Emmanuelle Schmitt and Yves Mechulam



**Supplementary Figure 1: Assembly of IC and Cryo-EM structure determination.**

(a) Cloverleaf representation *E. coli* initiator tRNA<sub>f<sup>Met</sup></sub> carrying the mutation C<sub>1</sub> A<sub>72</sub> into A<sub>1</sub> U<sub>72</sub> (right). (b) SDS-PAGE analysis of purification steps of the IC and cryo-EM image.

(b) Lane 1: Mixture of purified Pa-aIF2, Pa-aIF1 and Pa-aIF1A. Note that N-terminally tagged versions of Pa-aIF1, Pa-aIF2 $\alpha$  and Pa-aIF2 $\beta$  were used. aIF1 and aIF1A co-migrate. Lane 2: molecular weight marker (LMW, 97, 64, 43, 30, 20, 14 kDa, GE-Healthcare). Lane 3: 30S subunit from *P. abyssi* purified as described in Supplemental Materials and Methods. The band corresponding to phosphoenol pyruvate synthase (Pep synthase) is indicated by an arrow. Lane 4: IC complex after affinity column purification and concentration. Bands corresponding to aIF2 $\gamma$  and to aIF2 $\alpha$  are indicated. Bands corresponding to aIF1, aIF1A and aIF2 $\beta$  co-migrate with ribosomal proteins. The presence of tRNA was verified using staining with Ethidium Bromide.

(c) Cryo-EM image of IC collected on an FEI Cs-corrected Titan Krios microscope (-2.3  $\mu\text{m}$  defocus) operated at 300kV using a 4k x 4k FEI Falcon II camera (Netherlands Centre for Electron Nanoscopy, Leiden).

(d) Maximum-likelihood 3D classification scheme (see Online methods). The initiation factors bound to the 30S subunit (pink) are colored as follows: aIF2 green, aIF1A orange, aIF1 cyan and tRNA yellow.

(e) Gold-standard Fourier Shell Correlation (FSC) curves for post-processed maps. The resolutions of IC0-P<sub>OUT</sub> (in black) and IC1-P<sub>IN</sub> (in red) are 5.34  $\text{\AA}$ , and 7.5  $\text{\AA}$  respectively, using the 0.143 criterion.

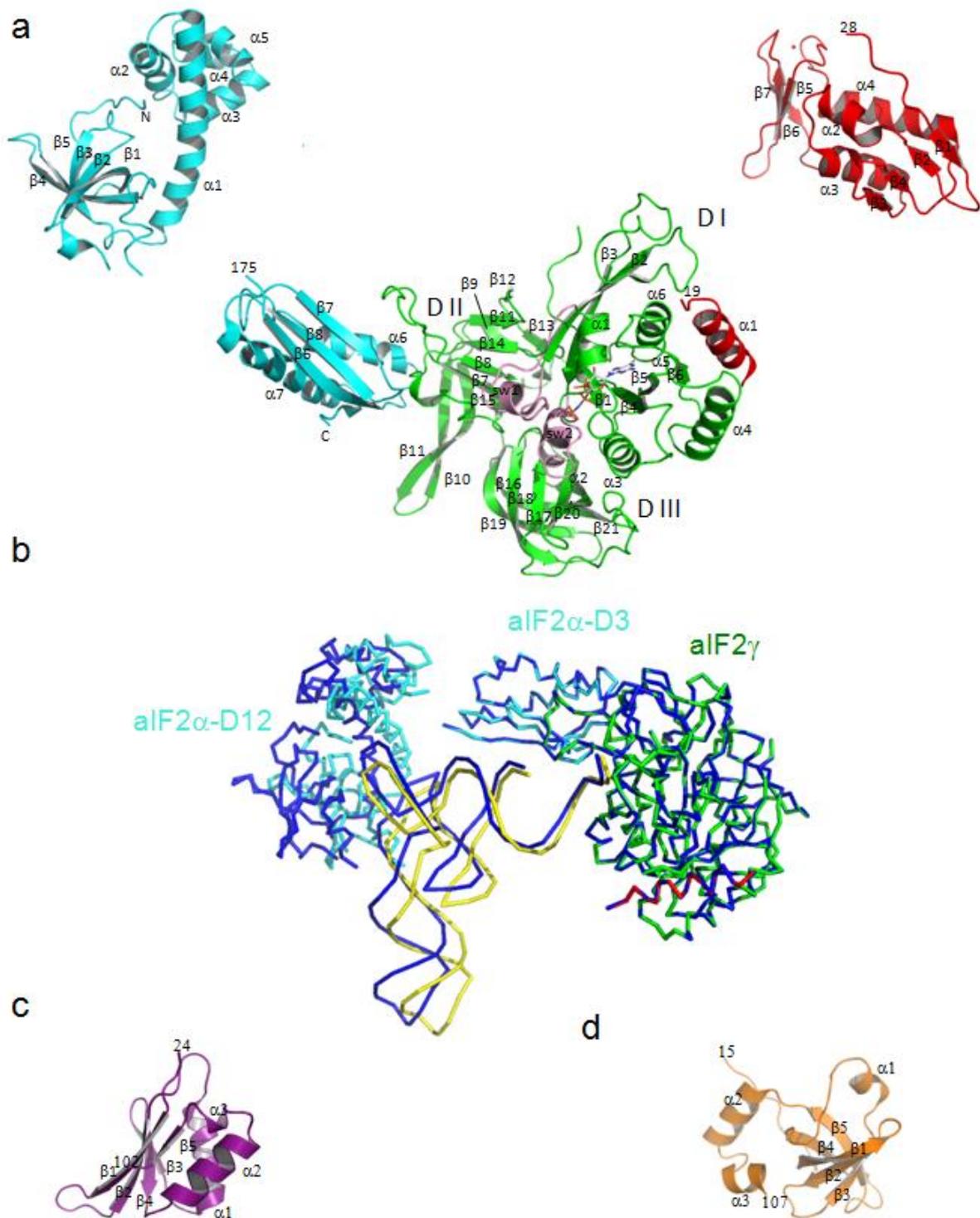
(f) Analysis of overfitting by cross-validation of the IC0-POUT complex model.

FSC<sub>work</sub> curve (green) corresponds to the refined model versus the half-map it was refined against. The FSC<sub>test</sub> curve (red) is calculated between the refined model and the other half-map. The black curve shows the FSC between the refined model and the full map used for refinement. The dashed line represents the highest resolution (5.8  $\text{\AA}$ ) used in refinements.

(g) Left: Surface of IC0-P<sub>OUT</sub> map colored according to local resolution (see Supplementary methods) using the color scale indicated on the right. Right: Same for IC1-P<sub>IN</sub>.

(h) The table shows the local resolutions in Å of the components of the initiation complexes as determined with RELION using a soft mask (with a 15-pixel soft-edge) surrounding the region of interest as described<sup>1</sup>. n.d. not determined.

The figures were drawn with Chimera<sup>2</sup> and Pymol<sup>3</sup>



**Supplementary Figure 2: Structures of initiation factors used for model building.**

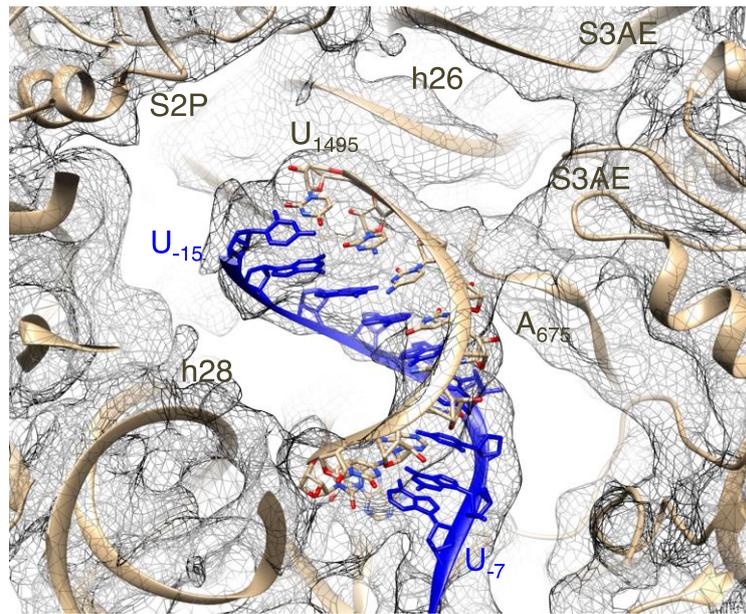
(a) aIF2 is composed of a core domain (aIF2 $\gamma$ -aIF2 $\alpha$ -D3-Nterminal helix of aIF2 $\beta$ ) onto which two mobile wings (aIF2 $\alpha$ -D12) and core domain of aIF2 $\beta$  are anchored<sup>4,5</sup> □ The structure of the core domain is shown on the center. The structure corresponds to a composite model made of Ss-aIF2 $\gamma$  (PDB entry 4RD4<sup>6</sup>) bound to  $\alpha$ 3 and the N-ter helix of  $\beta$  as observed in PDB entry 3V11<sup>7</sup>. aIF2 $\alpha$ -D12 is at the upper left, and the core of aIF2 $\beta$  is at the upper right. The subunits are colored as follows:  $\alpha$  in blue,  $\beta$  in red and  $\gamma$  in green, with its switch regions in pink.

(b) Superimposition of the crystal structure of the TC on the structure of the TC observed in IC0. The  $\gamma$  subunits of the two TC are superimposed. The TC from the crystal structure (PDB entry 3V11) is in dark blue. The TC from IC0 is colored as in Fig. 1a,  $\gamma$  in green,  $\alpha$  in cyan, the N-terminal helix of  $\beta$  in red and the tRNA in yellow. The view shows very good agreement of the two structures. Only aIF2 $\alpha$ -D12 domains are slightly rotated from their positions seen in the crystal structure, accompanied by a minor movement of the initiator tRNA position.

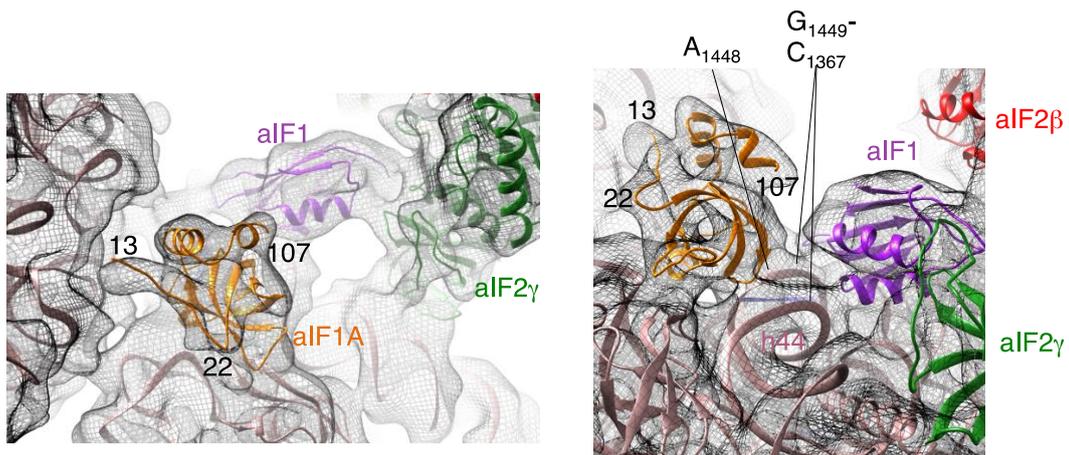
(c) Structure of aIF1 from *Methanocaldococcus jannashii* from PDB entry 4MO0.

(d) Structure of aIF1A from *Pyrococcus abyssi* from PDB entry 4MNO.

a



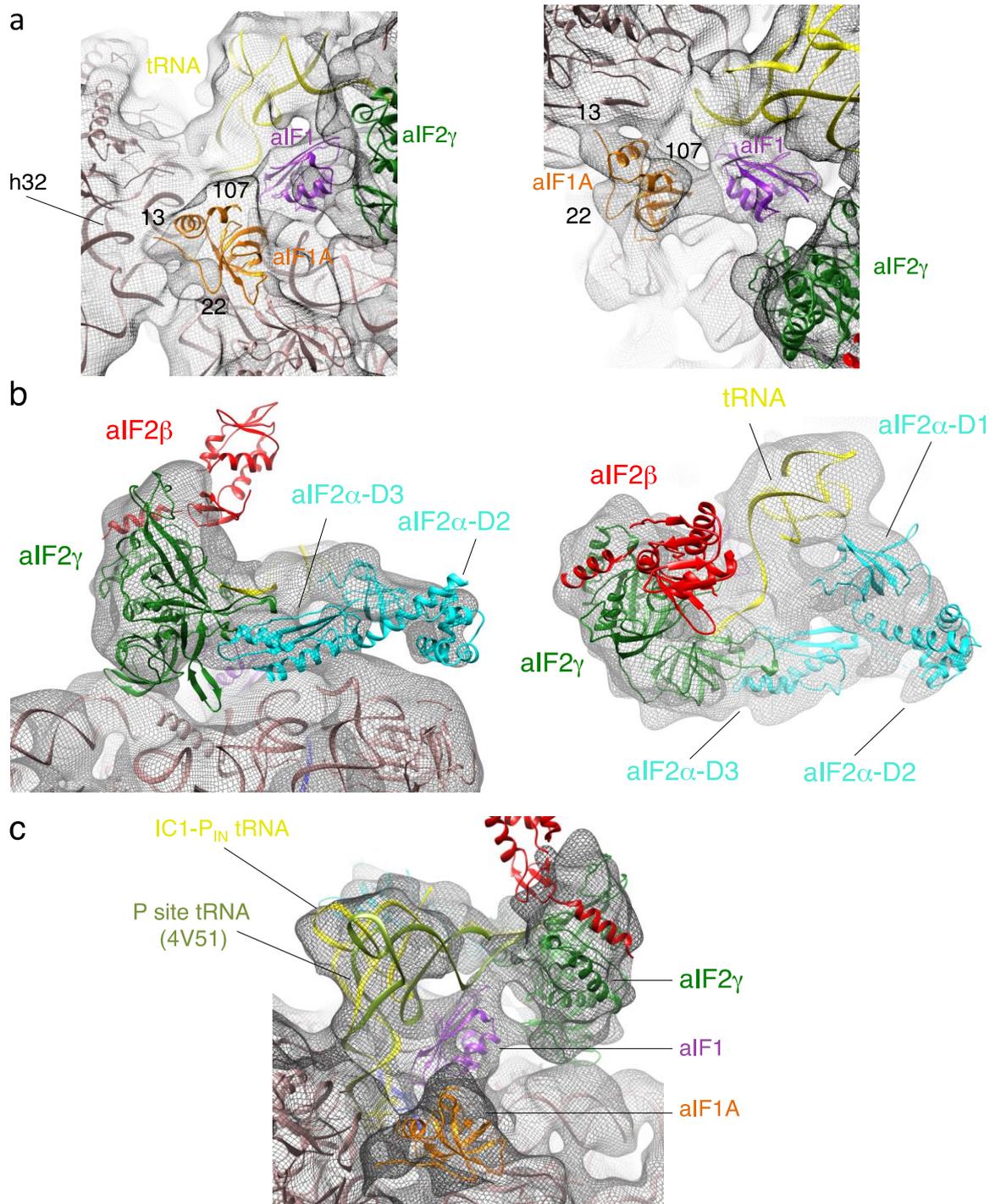
b



**Supplementary Figure 3:** Local fit of mRNA, aIF1 and aIF1A into IC0-P<sub>OUT</sub> map.

(a) Region of the mRNA containing the Shine-Dalgarno sequence base-paired with the rRNA as observed in the IC0-P<sub>OUT</sub> electron density map filtered at 5.3 Å resolution.

(b) Fitting of Pa-aIF1A (PDB entry 4MNO) and Mj-aIF1 (PDB entry 4MO0) in IC0-P<sub>OUT</sub> is shown in two different orientations. The cryo-EM map is filtered to 8.5 Å resolution.

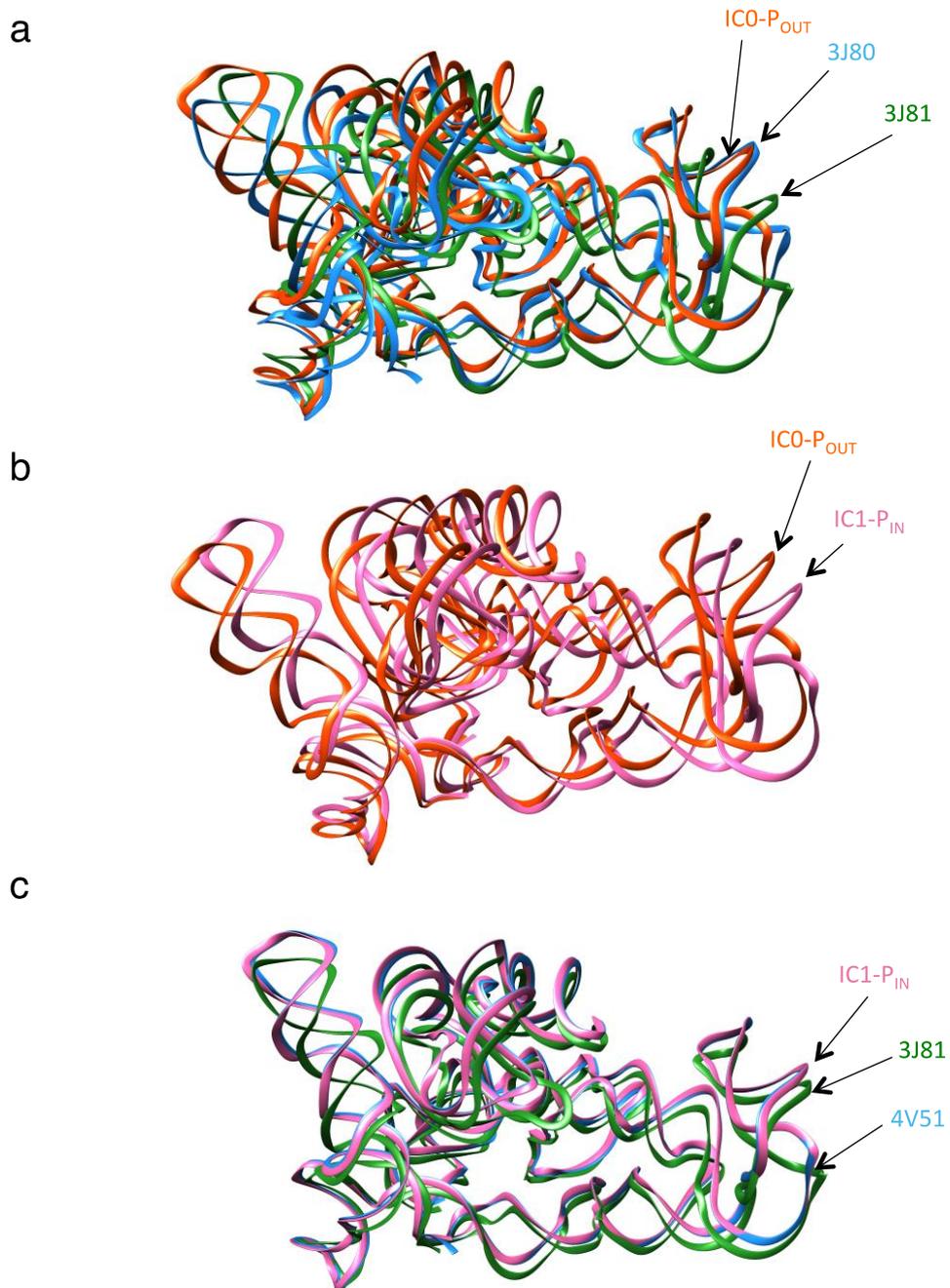


**Supplementary Figure 4:** Fitting of factors into IC1-P<sub>IN</sub> map.

(a) Fitting of Pa-aIF1A (PDB entry 4MNO) and Mj-aIF1 (PDB entry 4MO0) in IC1-P<sub>IN</sub>. The cryo-EM map is filtered to 12 Å resolution.

(b) Fitting of the ternary complex into IC1-P<sub>IN</sub> map.

(c) Fitting of initiator tRNA (yellow) within IC1-P<sub>IN</sub> map. The view shows that the structure of the P site initiator tRNA as observed in PDB entry 4V51<sup>8</sup> (light green) could not be readily rigid-body fitted in the electron density. A bending of the tRNA molecule was necessary to account for the density showing simultaneous binding of the initiator tRNA with the mRNA and aIF2.



**Supplementary Figure 5:** Head movement upon transition from IC0-P<sub>OUT</sub> to IC1-P<sub>IN</sub>.

Each panel compares position of the heads of the small ribosomal subunits after superimposition of the bodies.

(a) IC0-P<sub>OUT</sub> is compared to PDB entry 3J80 (40S:eIF1:eIF1A) and PDB entry 3J81 (py48S PIC)<sup>9</sup>.

(b) IC0-P<sub>OUT</sub> is compared to IC1-P<sub>IN</sub>.

(c) IC1-P<sub>IN</sub> is compared to PDB entry 3J81 (py48S PIC) and to PDB entry 4V51 (70S:mRNA:tRNA:paromomycin complex)<sup>8</sup>.

The figure shows that the orientation of the head in IC0-P<sub>OUT</sub> is similar to that in 40S:eIF1:eIF1A whereas that in IC1-P<sub>IN</sub> resembles those in py48S PIC and 70S:tRNA complex.

## REFERENCES

1. Nguyen, T.H. et al. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* **523**, 47-52 (2015).
2. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-12 (2004).
3. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. (2010).
4. Yatime, L., Mechulam, Y., Blanquet, S. & Schmitt, E. Structure of an archaeal heterotrimeric initiation factor 2 reveals a nucleotide state between the GTP and the GDP states. *Proc. Natl. Acad. Sci. U S A* **104**, 18445-50 (2007).
5. Stolboushkina, E. et al. Crystal structure of the intact archaeal translation initiation factor 2 demonstrates very high conformational flexibility in the alpha- and beta-subunits. *J. Mol. Biol.* **382**, 680-691 (2008).
6. Dubiez, E., Aleksandrov, A., Lazennec-Schurdevin, C., Mechulam, Y. & Schmitt, E. Identification of a second GTP-bound magnesium ion in archaeal initiation factor 2. *Nucleic Acids Res* **43**, 2946-2957 (2015).
7. Schmitt, E. et al. Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat. Struct. Mol. Biol.* **19**, 450-454 (2012).
8. Selmer, M. et al. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942 (2006).
9. Hussain, T. et al. Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159**, 597-607 (2014).

# 5 Discussion et perspectives

## 5.1 Discussion

Chez les eucaryotes, le facteur eIF1 joue un rôle primordial lors du démarrage de la traduction. Cependant, le rôle de son orthologue aIF1 chez les archées restait très peu connu et l'étude de sa fonction lors du démarrage de la traduction archée restait à mettre en évidence.

Chez les eucaryotes, eIF1 se lie à la sous-unité du ribosome (Lomakin et al. 2003) même en l'absence des autres facteurs (Maag et Lorsch 2003).

Les études par anisotropie de fluorescence nous ont permis de mettre en évidence la liaison spécifique d'aIF1 à la sous-unité 30S. Le Kd d'aIF1 pour la 30S est très petit (de l'ordre du nanomolaire) et, tout comme pour le facteur orthologue eucaryote, la liaison d'aIF1 ne nécessite pas la présence des autres facteurs de démarrage dans le milieu. Par ailleurs, nous avons pu mettre en évidence la liaison coopérative d'aIF1 et de l'ARNm à la 30S. En effet, l'affinité d'aIF1 pour la 30S est augmentée (d'au moins un facteur 5) lorsque cette dernière est déjà en complexe avec l'ARNm.

De façon intéressante, la modélisation d'un ARNt au site P du ribosome dans la structure cristallographique de Tt-eIF1:40S prédit qu'eIF1 empêche la liaison définitive de ce dernier *via* la présence de la boucle de résidus basiques d'eIF1 qui entrainerait un choc stérique avec la tige anticodon de l'ARNt<sup>Met</sup> (Rabl et al. 2011). De plus, cette boucle basique pourrait être en interaction avec l'ARNr 18S, et l'appariement entre l'anticodon de l'ARNt<sup>Met</sup> et le codon AUG de l'ARNm pourrait perturber l'interaction entre la boucle basique et la 40S, favorisant ainsi la dissociation d'eIF1. Des études d'eIF1 de levure ont permis de mettre en évidence l'implication de trois résidus basiques de cette boucle dans la fonction d'eIF1 *in vivo* et dans la liaison à la 40S *in vitro* (Martin-Marcos et al. 2013). Ces résidus sont impliqués dans le recrutement du complexe ternaire et permettent de prévenir la dissociation prématurée d'eIF1, pouvant entrainer un changement de conformation du ribosome passant alors prématurément en position fermée sur l'ARNm sur un codon pouvant ne pas être un codon AUG.

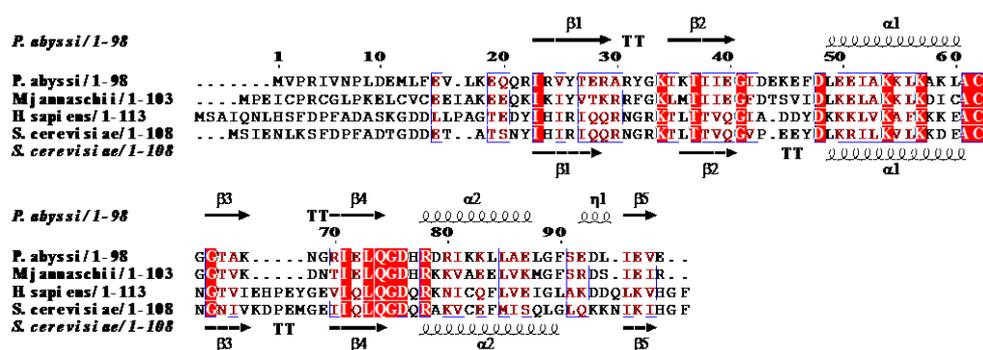


Figure 5.1 : Alignement des séquences des facteurs de démarrage aIF1 archées et eIF1 eucaryotes

L'alignement de la structure cristallographique de Mj-aIF1 sur celle de Tt-eIF1:40S nous a permis de constater que la boucle de résidus basiques présente chez le facteur archée aIF1 pouvait de la même façon être cruciale dans le rôle lors de la formation du complexe de démarrage des archées (voir figure 3.15 partie anisotropie). Cependant, même si le caractère basique de la boucle est conservé chez les facteurs e/aIF1, les résidus ne sont pas strictement identiques entre les facteurs aIF1 archées et eIF1 eucaryotes. Nous avons donc choisi de muter les résidus 32, 33 et 35 de cette boucle. Nous avons pu déterminer par anisotropie de fluorescence les Kd de ces variants. Nous montrons que deux résidus sont impliqués dans la liaison d'aIF1 à la 30S : la tyrosine 33 (Y33), et la lysine 35 (K35). Par ailleurs, ces résidus ne semblent pas impliqués de façon similaire dans la liaison d'aIF1 à la 30S. En effet, le résidu Y33 est impliqué dans la stabilisation de l'ARNm sur la 30S, tandis que le résidu K35 serait plutôt orienté du côté de la liaison d'aIF1 à la 30S (cf tableau 1).

Résidu étudié	Implication d'aIF1 dans la liaison au complexe 30S:ARNm	Liaison coopérative de l'ARNm et d'aIF1 à la 30S	Fonction dans la stabilisation de la 30S sur l'ARNm	Fonction dans la déstabilisation du CI sur l'ARNm
WT		Oui		
R32	Très faible	Non testée	Non	Oui
Y33	Oui	Non	Oui ++	Oui
K35	Très impliqué	Oui	Oui +	
$\Delta N$	Non	Non testée	Non	Oui

Tableau 1

Ensuite, nous avons utilisé la méthode du toeprint afin d'observer la formation et la stabilité du ribosome en présence de différents facteurs de démarrage sur un ARNm d'intérêt, et ainsi étudier le rôle d'aIF1 dans la formation du complexe de démarrage. Comme cela était attendu d'après les résultats obtenus par anisotropie de fluorescence, la présence du facteur aIF1 sur la 30S suffit à stabiliser cette dernière sur l'ARNm. De plus, nous avons pu confirmer l'implication des résidus Y33 et K35 dans la stabilisation de la 30S sur l'ARNm. En effet, la stabilisation de la 30S n'est effective qu'à 23% et 69% en absence de l'un ou l'autre de ces résidus (cf tableau 2).

Variants d'aIF1 utilisé	Kd d'aIF1 au complexe 30S:ARNm (en nM)(*)	Kd d'aIF1 à la 30S (en nM)(*)	signal relatif +12-13-14 (aIF1 seul)	signal relatif +14-15 (dans le CI)
Sans aIF1				100% ± 2
aIF1WT	0.5 ± 0.2	25 ± 4	100% ± 8	52% ± 5
aIF1-RA32	2.4 ± 0.5		100% ± 3	84% ± 3
aIF1-YA33	11.7 ± 2.8	14 ± 1	23% ± 12	93% ± 16
aIF1-KA35	45 ± 8	381 ± 46	69% ± 15	
aIF1-RA32-YA33	61 ± 15		21% ± 1	100% ± 7
aIF1-RA32-KA35			77% ± 2	81%
aIF1-ΔN	0.7 ± 0.1		107% ± 4	98% ± 8

Tableau 2

(\*) déterminé par anisotropie de fluorescence

En rouge : les valeurs se différenciant nettement des valeurs obtenues avec aIF1 WT.

Stabilisation : % signal +12-13-14 =  $(\sum_{12/13/14} \text{aIF1mutant} / \sum_{12/13/14} \text{aIF1WT}) \times 100$

Déstabilisation : % signal +14-15 =  $(\sum_{14/15} \text{CI} \times 100) / \sum_{14/15} \text{CI sans aIF1}$ . Le 100% du signal en position +14 +15 est défini avec la moyenne des expériences réalisées avec le CI sans aIF1.

Cependant, la stabilisation de la 30S sur l'ARNm ne dépend pas de la position 1 ou 2 d'un codon de démarrage AUG sur l'ARNm, et même ne nécessite pas la présence d'un codon AUG. A la différence des ARNm eucaryotes, les ARNm archées possèdent fréquemment une séquence SD s'appariant avec l'ARNr 16S de la 30S. Nous avons montré que la stabilisation du complexe aIF1:30S sur l'ARNm dépendait exclusivement de la présence d'une séquence SD en amont du codon de démarrage. Cette séquence doit posséder au moins la séquence consensus  $_{(-13)}\text{GGAGG}_{(-9)}$  suivie des bases G<sub>-7</sub> et A<sub>-6</sub> pour stabiliser le complexe et provoquer l'arrêt de la RT dans une expérience de toeprint.

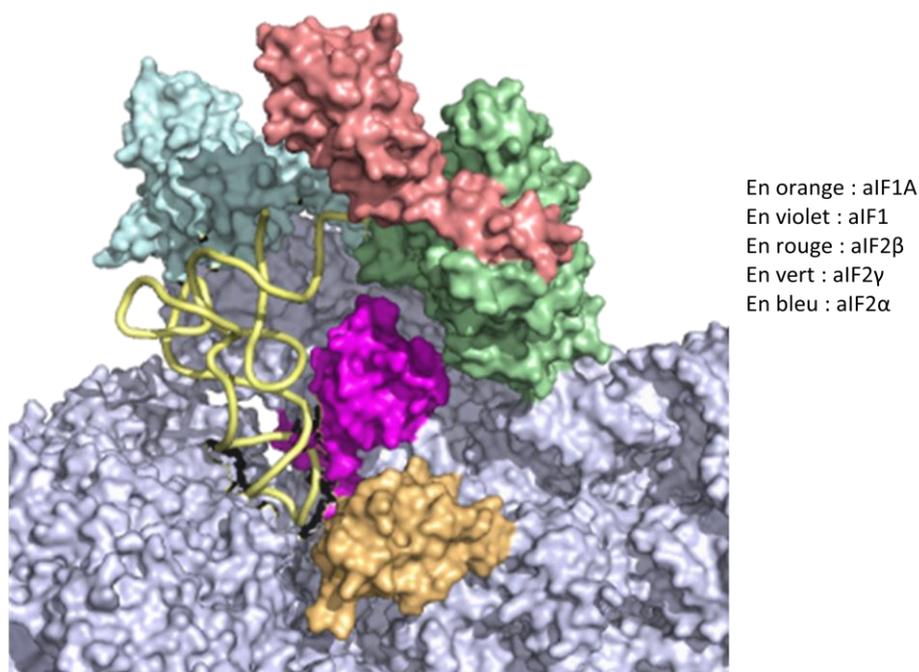


Figure 5.2 : Modèle en surface moléculaire du complexe IC1-P<sub>in</sub> par Cryo-EM

Nous avons ensuite étudié le rôle d'aIF1 dans la formation du complexe de démarrage. Tout comme pour le facteur eucaryote (Martin-Marcos et al. 2013), le facteur aIF1 déstabilise le complexe de démarrage (CI) centré sur le codon AUG. En effet, le signal observé en position +14 et +15 (reflétant une conformation du complexe de démarrage dans lequel l'ARNt<sup>Met</sup> est définitivement apparié au codon AUG au niveau du site P du ribosome), n'est que de 52% en présence d'aIF1. Chez les eucaryotes, eIF1 stabilise le complexe dans une conformation ouverte, permettant le balayage de l'ARNm. Chez les archées, aIF1 pourrait de même favoriser une conformation ouverte du ribosome dans laquelle l'ARNt ne serait pas définitivement appariée au codon de l'ARNm, empêchant ainsi la formation prématurée d'un mauvais complexe de démarrage. Cette hypothèse est appuyée par la structure IC1-P<sub>in</sub> (voir figure 1 ci-dessus). Dans cette structure, l'anticodon de l'ARNt est apparié au codon AUG de l'ARNm. aIF1 est toutefois encore présent. L'atténuation de la densité électronique d'aIF1 dans le complexe IC1-P<sub>in</sub> reflète l'aspect dynamique de cette conformation. L'accommodation définitive de l'ARNt pourrait être gênée par la présence d'aIF1 sur la 30S. Ainsi, la présence d'aIF1 viendrait déstabiliser le complexe de démarrage sur l'ARNm dans une conformation dans laquelle l'ARNt ne serait pas définitivement stabilisé au site P du ribosome. De plus, tout comme chez les eucaryotes (Pestova, Borukhov, et Hellen 1998), le facteur aIF1A renforcerait l'effet déstabilisateur d'aIF1 du complexe de démarrage centré sur le codon AUG. aIF1 et aIF1A auraient donc aussi une action synergique.

L'étude des résidus de la boucle basique d'aIF1 par toeprint dans la formation du complexe de démarrage (CI), et donc dans la fonction déstabilisatrice du complexe de démarrage centré sur le codon AUG par aIF1 nous a permis de mettre en évidence l'implication de 2 résidus de cette boucle. Le résidu Y33 vu précédemment comme étant aussi impliqué dans la liaison d'aIF1 à la 30S et dans la stabilisation du complexe aIF1:30S, mais aussi le résidu arginine 32 (R32). Les pourcentages de signaux en position +14 et +15 obtenus lorsque ces résidus d'aIF1 sont échangés en alanine dans le CI sont respectivement de 84% et 93% lors de la mutation des résidus 32 et 33. Ces résultats reflètent l'incapacité d'aIF1 à assurer sa fonction déstabilisatrice du complexe de démarrage centré sur le codon AUG lorsque l'un de ces résidus est muté en alanine. L'implication du résidu K35 reste à déterminer. D'autre part, le domaine N-terminal du facteur aIF1 est aussi impliqué dans cette fonction. Le pourcentage de signal obtenu en positions +14 et +15 est maintenu à 98%. Ce domaine N-terminal est intéressant quant à la fonction du facteur aIF1. Celui-ci n'est pas impliqué dans la liaison d'aIF1 à la 30S ni dans la fonction stabilisatrice d'aIF1 de la 30S sur l'ARNm, mais est très fortement impliqué dans la fonction déstabilisatrice d'aIF1 du complexe de démarrage centré sur le codon AUG. Dans les structures des complexes de démarrage archées, IC0-Premote et IC1-Pin, le domaine N-terminal d'aIF1 est en contact avec la sous-unité gamma d'aIF2. Compte-tenu des résultats obtenus en toeprint, on peut proposer que ces contacts participent à l'action déstabilisatrice d'aIF1.

Pour finir, il a été montré que le facteur eIF1 est essentiel pour empêcher le démarrage de la traduction sur un codon AUG dans un mauvais contexte (Yoon et Donahue, 1992 ; Pestova et Kolupaeva, 2002 ; Valasek, 2004). Nous avons ainsi voulu tester la capacité d'aIF1 à empêcher la formation du complexe de démarrage centré sur le codon AUG dans le cas où un ARNt élongateur serait utilisé à la place d'un ARNt initiateur. Dans ce cas, la présence d'aIF1 dans le complexe de démarrage permet de déstabiliser davantage le complexe centré sur le codon de démarrage lorsqu'un ARNt élongateur est présent dans le milieu (voir tableau 3).

ARNt utilisé	signal relatif +14-15
Met-ARNt <sub>i</sub> <sup>Met</sup> A <sub>1</sub> -U <sub>72</sub>	52% ± 5
Met-ARNt <sub>m</sub> <sup>Met</sup>	27% ± 6
Met-ARNt <sub>m</sub> <sup>Met</sup> 3G3C	62% ± 3

Tableau 3

aIF1 pourrait donc faciliter l'accommodation des ARNt initiateurs. Cet effet pourrait être relié à la particularité de la tige anticodon des ARNt initiateurs. En effet, dans les structures de complexe de démarrage archée, nous observons que la boucle basique d'aIF1 se situe juste devant les trois paires GC de l'ARNt initiateur. Même si la résolution limitée de ces complexes ne permet pas d'identifier les détails moléculaires de l'interaction entre aIF1 et l'ARNt initiateur au site P, la localisation observée pour aIF1 est en accord avec une possible interaction avec les trois paires GC. aIF1 pourrait aussi être sensible à une conformation particulière de l'ARNt<sub>i</sub> provoquée par la présence des trois paires GC (Barraud et al. 2008). En effet, lorsque les trois paires de bases GC caractéristiques des ARNt initiateurs sont insérées dans la tige anticodon d'un ARNt élongateur, l'effet déstabilisateur d'aIF1 du complexe de démarrage redevient identique à celui observé lorsqu'un ARNt initiateur est présent dans le milieu (tableau 3).

Finalement, lorsque le complexe de démarrage se trouve dans un contexte favorable au démarrage de la traduction, le facteur aIF1 se dissocierait de la 30S, permettant ainsi l'accommodation définitive de l'ARNt initiateur au site P du ribosome. Cette dernière conformation peut être observée dans la structure d'un complexe également étudié par Cryo-EM (mais pas encore publié), dans laquelle le facteur aIF1 est absent du complexe de démarrage et dans laquelle l'ARNt initiateur est définitivement stabilisé au site P du ribosome (complexe nommé IC2-P<sub>in</sub>). De manière cohérente, cette conformation peut être observée par toeprint lorsqu'aIF1 n'est pas présent dans le milieu, il n'y a plus de déstabilisation du complexe de démarrage centré sur le codon AUG et le signal est restreint aux positions +14 et +15.

## 5.2 Perspectives

Durant mes trois années de thèse, l'obtention des différents acteurs du démarrage de la traduction archée nous a permis d'obtenir la structure par Cryo-EM du complexe de démarrage archée. De plus, nous avons développé au laboratoire des outils d'étude biochimique par anisotropie de fluorescence et par toeprint. Ces résultats nous ont permis d'enrichir nos connaissances sur le démarrage de la traduction chez les archées, entre autres sur la fonction du facteur aIF1. Malgré tout, la compréhension des mécanismes de formation de ces complexes peut encore être approfondie en utilisant par des études biochimiques et structurales. En particulier, la mise en place des méthodes de toeprint et d'anisotropie ouvre la voie à l'étude des différents facteurs de démarrage.

Si nous arrivons à marquer avec un fluorophore les facteurs aIF1A et aIF2 de façon convenable, nous pourrions étudier l'action synergique des différents facteurs lors de la liaison à la 30S par anisotropie de fluorescence. Le marquage effectif de ces facteurs pourra par ailleurs nous permettre de mettre au point des expériences de stopped-flow, permettant l'étude à l'état pré-stationnaire des complexes de démarrage.

Nous avons étudié par anisotropie de fluorescence la liaison du facteur aIF1 à la 30S seule ou complexée à l'ARNm. Cependant, comme nous avons pu le constater par toeprint, les facteurs aIF1 et aIF1A semblent agir de façon synergique lors de la formation du complexe 30S:ARNm. Nous avons donc commencé à étudier la formation du complexe aIF1:aIF1A:30S:ARNm par anisotropie de fluorescence. Aucun effet sur le titrage d'aIF1 par le complexe 30S:ARNm n'a pu être constaté lors de l'ajout du facteur aIF1A dans le milieu. Ce résultat peut être dû au fait que le  $K_d$  d'aIF1 pour le complexe 30S:ARNm est trop faible compte tenu des concentrations utilisées lors de nos expériences. Même si l'ajout du facteur aIF1A pourrait avoir comme conséquence de diminuer le  $K_d$ , cet effet ne serait alors pas observable. Toutefois, nous avons pu déterminer un  $K_d$  plus élevé d'aIF1 pour la 30S. un titrage d'aIF1 avec le complexe aIF1A:30S en l'absence d'ARNm pourrait alors nous permettre de constater un possible effet du facteur aIF1A sur la liaison d'aIF1 à la 30S. D'autre part, il serait intéressant de poursuivre les titrages d'aIF1 par anisotropie de fluorescence avec le complexe de démarrage total, afin de pouvoir constater un éventuel effet sur la liaison d'aIF1 à la 30S.

En parallèle, la caractérisation des résidus du facteur aIF1A impliqués dans la fonction de ce dernier dans le démarrage de la traduction pourra être réalisée par toeprint. Différentes versions d'aIF1A tronqué dans son domaine N-terminal ont été utilisées pour ces études. En effet, le domaine N-terminal d'aIF1A pourrait interagir avec le facteur aIF1, et nous avons cherché à mettre son influence en évidence. Jusqu'à ce jour, aucun effet notable de ce domaine n'a pu être constaté. La caractérisation des résidus d'aIF1A pouvant être impliqués dans la formation du complexe de démarrage par toeprint reste à mettre en place.

Pour finir, les avancées au laboratoire concernant le facteur eIF2 eucaryote m'ont permis de tenter d'exprimer le facteur eIF2 humain chez *E. coli*. En effet, jusqu'à il y a peu de temps, l'expression du facteur eIF2 de levure total ne fonctionnait pas chez *E. coli*. La sous-unité gamma restait insoluble. La découverte de l'implication de la protéine chaperonne CDC123 dans l'assemblage des sous-unités  $\alpha$  et  $\gamma$  d'eIF2 (Panvert et al. 2015) est un point de départ pour l'expression de ce facteur pour d'autres espèces eucaryotes, CDC123 étant une protéine très conservée. J'ai déjà réalisé différents tests d'expression des sous-unités d'eIF2 humain dans différentes conditions (variation du milieu de culture, variation de la température de culture ou d'expression, variation du temps d'expression,

variation des plasmides utilisés). A ce jour, seule la sous-unité  $\alpha$  a pu être surexprimée. La sous-unité gamma a pu être isolée soluble, mais dans des quantités ne permettant pas sa purification en grande quantité, et les différentes tentatives d'expression en grande quantité n'ont rien donné. D'autres tentatives d'expression et de purification de ce facteur eIF2 humain pourront donc être entreprises afin de pouvoir débiter des études structurales et fonctionnelles de ce facteur humain.



# 6 Matériels et méthodes

## 6.1 Inactivation de gènes chez E.coli par la technique lambda red

### 6.1.1 Souches bactériennes et plasmides

L'inactivation des gènes d'ARN<sub>t</sub><sup>initiateurs</sup> est réalisée dans la souche MG1655 d'*Escherichia coli*. Cette souche dérive de la souche W1485 dérivant elle-même de la souche sauvage K-12 et a été conservée en laboratoire avec le moins de manipulations génétiques possible.

Génotype : (F- lambda- ilvG- rfb-50 rph-1)

Sérotype : OR : H48 : K-

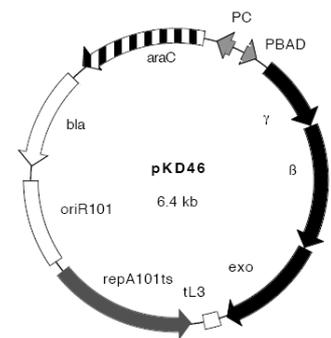
Toutes les délétions primaires (inactivation du ou des premiers gènes codant(s) pour un ARN<sub>t</sub><sup>initiateur</sup>) ou secondaires (inactivation du ou des second gènes codant(s) pour les ARN<sub>t</sub><sup>initiateurs</sup>) sont effectuées dans la souche MG1655.

Le plasmide contenant l'opéron Red ( $\gamma$   $\beta$   $\text{exo}$ ) est le plasmide pKD46 (6329pb). Il est à faible nombre de copies et thermosensible à 37°C. Ce plasmide dérive de pDK20.

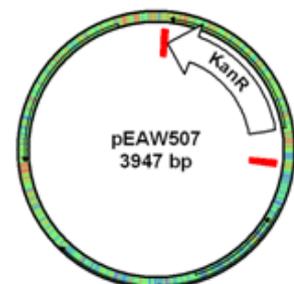
Il possède les gènes  $\gamma$ ,  $\beta$  et  $\text{exo}$  placés sous le contrôle du promoteur pBAD inductible à l'arabinose, le gène  $\text{araC}$ , le gène  $\text{bla}$  de résistance à l'ampicilline (ou le gène de résistance à la gentamicine), une séquence tL3 (qui code pour une petite ORF responsable du phénotype d'inhibition de l'hôte et qui permet d'améliorer le nombre de recombinant) en aval du site de terminaison du gène  $\text{exo}$ , et le gène codant la protéine RepA101ts qui initie la réplication à 30°C mais pas à 37°C.

pKD46 entraîne un meilleur taux de recombinaison que le plasmide pKD20.

Les souches transformées avec ces plasmides sont sélectionnées sur milieu LB amp.



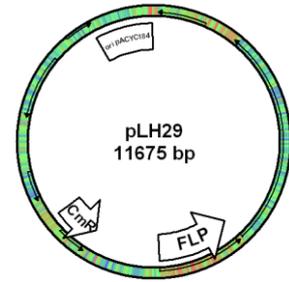
Le plasmide permettant d'obtenir par PCR le fragment d'ADN contenant le gène de résistance à la kanamycine (KanR) flanqué de deux régions FRT ainsi que les régions homologues au gène à inactiver est pEAW507. Les souches transformées avec ce plasmide ou les produits de PCR provenant de ce plasmide sont sélectionnées sur milieu LB kan.



Le plasmide permettant d'exprimer la protéine FLP est le plasmide faible nombre de copies pLH29 de 11675 pb. Il contient entre autres le gène de résistance au

chloramphénicol, le gène *lacI*, le promoteur du gène codant pour la protéine FLP inductible à l'IPTG et le gène codant pour la recombinaise FLP.

Les souches transformées avec ces plasmides sont sélectionnées sur milieu LB cam.



Les plasmides utilisés pour les tests de complémentation sont :

- pBSTnav2 : il contient le gène de résistance à l'ampicilline (*bla*)
- pBSTnav2 ARN<sub>t</sub> fmet : contient le gène *bla* et le gène codant pour l'ARN<sub>t</sub><sup>initiateur</sup> de *E. coli* non muté
- pBSTnav2 ARN<sub>t</sub> sulfolobus : contient le gène *bla* et le gène codant pour l'ARN<sub>t</sub><sup>initiateur</sup> de l'archée *Sulfolobus solfataricus*
- pBSTnav2 ARN<sub>t</sub> A1U72 : contient le gène *bla* et le gène codant pour l'un des ARN<sub>t</sub><sup>initiateur</sup> de *E. coli* muté dans ses bases 1 et 72 (C1 → A1 / A72 → U72).

Les souches transformées avec ces plasmides sont sélectionnées sur milieu LB amp.

### 6.1.2 Préparation des produits de PCR possédant le gène KanR, 2 séquences FRT et 2 régions homologues au(x) gène(s) à inactiver

Les oligonucléotides (oligos) utilisés pour la PCR doivent avoir 50 à 70 nucléotides d'homologie avec les séquences encadrant la séquence génomique à éliminer.

Les séquences en 3' des oligos permettent d'amplifier le gène KanR du plasmide pEAW507.

Voir ci-dessous les séquences des oligos choisis pour l'amplification du gène KanR en vu d'inactiver le gène *metY* :

MetYdelfor : GCAGTATTTGCATTTTTTACCCAAAACGAGTAGAATTTGCCACGTTTCAGGCGTAGAGGATCCCCG  
GAAGT

MetYdelback : CGAAGGCCGAAGTCTTCACAGTATATTTGAAAAAGGACTCTAAGGGAAAGCCCTTTCGTCTTC  
AAGAAT

Séquence des oligos choisis pour l'amplification du gène KanR en vu d'inactiver les gènes *metZ*, *W*, *V* :

MetZWVdelfor : AAAAAAAGGTTGCATGAAAACGCGAGCGGAGTATAGTGCGCATCCACGGACGTAGAGGAT  
CCCGGAAGT

MetZWVdelback : TTTATGGGAGACCAGAAACAAAAAACACCCGTTAGGGTGTTCAATAATCCCTTTCGTCT  
TCAAGAAT

Légende : homologie avec le locus MetY

homologie avec les locus metZ ou metV

homologie avec le locus FRT du plasmide pEAW507

Avant la PCR, le plasmide pEAW507 est linéarisé avec l'enzyme de restriction HincII.

La PCR est ensuite réalisée avec le plasmide pEAW507 linéarisé et les oligos décrits ci-dessus avec une température d'hybridation de 60°C (30 cycles).

Les produits PCR (environ 1,2 kb) sont digérés par l'enzyme de restriction DpnI puis précipités à l'éthanol et resuspendus dans 5µl d'eau.

Un contrôle de la digestion du plasmide pEAW507 par l'enzyme HincIII, ainsi qu'un contrôle des produits de PCR obtenus sont réalisés par migration sur gel agarose 1%.

Les produits issus de la PCR sont appelés PCRmetY et PCRmetZWV et permettent respectivement l'inactivation des gènes metY et metZWV.

### **6.1.3 Inactivation du gène metY (inactivation primaire)**

Les souches MG1655 + pKD46 compétentes sont transformées par électroporation avec le produit de PCR PCRmetY, puis repressent dans 1mL de milieu LB + L-arabinose et mises en culture liquide 2 heures à 30°C puis au moins 4 heures à 37°C afin de permettre la perte du plasmide pKD46. Par la suite, les cellules sont étalées sur des boîtes de Petri contenant de la kanamycine et mises en culture à 37°C sur la nuit.

Le lendemain, les clones possiblement positifs sont restriés sur boîte LB kan, puis une PCR sur colonie est réalisée afin de contrôler l'insertion du gène KanR dans la région du gène à inactiver (température d'hybridation : 60°C, 55°C ou 51°C).

Oligonucléotides utilisés pour la PCR de vérification de l'inactivation du gène metZWV :

verZfor : GGCGCAACGAAGATAACAACCGCCGC

verZback : GTATTTTGCATGGGGCGACGTGGCGA

Ces oligonucléotides se situent en amont ou en aval des gènes metZWV.

Les produits de cette PCR sont déposés sur un gel agarose 1%.

Un contrôle de la perte du plasmide pKD46 est réalisé en repiquant les colonies sur boîte de Pétri contenant de l'ampicilline et de la kanamycine. Ces colonies sont ensuite cultivées à 30°C ou 37°C sur la nuit.

Les clones positifs (c'est-à-dire ayant perdu le plasmide pKD46 et étant inactivés dans le/les gènes choisis) sont sélectionnés pour la suite des manipulations et les souches résultantes sont conservées à -80°C dans du glycérol 30%.

#### **6.1.4 Élimination du marqueur Kan puis élimination du plasmide pLH29**

Les souches sélectionnées sont rendues électrocompétentes puis transformées avec le plasmide pLH29. Les transformants sont alors mis en culture sur boîte contenant un milieu LB cam à 37°C sur la nuit.

Le lendemain, une colonie est inoculée (Pal $\Delta$ metZWV) dans 2mL de milieu LB cam contenant 1mM d'IPTG puis mise en culture à 37°C sur la journée et à température ambiante pendant 2 jours.

Les cellules sont par la suite étalées sur boîtes LB cam, puis 100 colonies sont repiquées sur boîte LB kan cam et LB cam.

Les clones poussant seulement sur LB cam et pas sur LB kan sont sélectionnés et la perte du marqueur KanR est confirmée grâce à une PCR sur colonie en utilisant les couples d'amorces VerZWVfor/VerZWVback (voir ci-dessus).

Le plasmide pLH29 est par la suite curé en ensemençant 5mL de milieu LB avec les clones positifs ayant perdu le marqueur KanR, puis ces clones sont mis en culture sur la nuit à 37°C.

Le lendemain, un étalement des cultures sur boîte LB cam ou LB est réalisé afin de constater ou non la perte du plasmide pLH29.

Si des clones poussent sur la boîte LB cam, cela indique qu'ils n'ont pas encore perdu pLH29 et l'opération de culture en milieu LB est répétée jusqu'à la perte effective du plasmide.

Lorsque le plasmide est perdu, la délétion primaire est obtenue ( $\Delta$ metZWV).

#### **6.1.5 Test de complémentation fonctionnel**

Des tests de complémentation fonctionnels sont réalisés afin de constater que les changements phénotypiques observés (voir partie résultats et discussion) sont dûs à la perte du gène d'intérêt. De

plus, le choix du gène codant pour un ARN<sub>t</sub> initiateur qui pourra compléter la 2<sup>ème</sup> délétion est aussi nécessaire.

Les souches MG1655 compétentes et inactivées dans les gènes metZWV (inactivation primaire) sauvages mais portant toujours la cassette KanR ou des souches MG1655 sont transformées ou non avec différents plasmides :

- pBSTnav2 ARN<sub>t</sub> Sulfolobus
- pBSTnav2 ARN<sub>t</sub> fmet
- pBSTnav2 ARN<sub>t</sub> A1U72
- pBSTnav2

Les cellules compétentes sont donc transformées, étalées sur boîte contenant de l'ampicilline ou de la kanamycine, puis repiquées le lendemain.

La taille des colonies peut alors être observée et selon l'observation phénotypique d'une complémentation, un ARN<sub>t</sub> initiateur sera choisi pour compléter les cellules au moment de l'inactivation secondaire de leur ARN<sub>t</sub> initiateur.

#### **6.1.6 Inactivation du gène metY (inactivation secondaire)**

Les souches PalΔMetZWV complémentées avec le plasmide contenant le gène de l'ARN<sub>t</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> (pBSTnav2A1-U72) compétentes sont transformées par électroporation avec le plasmide pKD46 contenant le gène de résistance à la gentamicine et le produit de PCR PCRmetY. Les cellules sont reprises dans 1mL de milieu LB + L-arabinose et mises en culture liquide 2 heures à 30°C puis au moins 4 heures à 37°C afin de permettre la perte du plasmide pKD46. Par la suite, les cellules sont étalées sur des boîtes de Petri contenant de la kanamycine et de l'ampicilline et mises en culture à 37°C sur la nuit.

Le lendemain, les clones possiblement positifs sont restriés sur boîte LB kan+amp, puis une PCR sur colonie est réalisée afin de contrôler l'insertion du gène KanR dans la région du gène à inactiver (température d'hybridation : 60°C, 55°C ou 51°C).

Oligonucléotides utilisés pour la PCR de vérification de l'inactivation du gène metY :

verYfor : AGTATTCAACAAATGAAAGTGAAGTGG

verYback : TAACAAAAACCCCGATAAATCGGGGC

Ces oligonucléotides se situent en amont ou en aval du gène metY

Les clones contrôlés positifs déletés de leur gène metY et ayant leur gène metY remplacés par la cassette de résistance à la kanamicine, et possédant le plasmide pBSTnav2A1-U72, nommés Pal $\Delta$ metZWV $\Delta$ MetY:KanR+A1U72 sont conservés à -80°C.

## **6.2 Purification des sous-unités 30S**

Tampon A : 0.1M NH<sub>4</sub>Cl ; 10.5mM MgAc ; 6mM  $\beta$ -M ; 0.1mM PMSF ; 20mM HEPES pH 7.5 ; 0.1mM benzamidine ; 0.5mM EDTA

Tampon de dialyse : 0.1M NH<sub>4</sub>Cl ; 2.5mM MgAc ; 6mM  $\beta$ -M ; 0.1mM PMSF ; 20mM HEPES pH 7.5 ; 0.1mM benzamidine ; 0.5mM EDTA

Tampon coussin de saccharose : 1M NH<sub>4</sub>Cl ; 10.5mM MgAc ; 1.1M saccharose ; 20mM HEPES pH 7.5 ; 0.1mM benzamidine ; 25 $\mu$ M PMSF

Tampon de saccharose 10% : 0.1M NH<sub>4</sub>Cl ; 2.5mM MgAc ; 6mM  $\beta$ -M ; 20mM HEPES pH 7.5 ; 0.1mM benzamidine ; 0.5mM EDTA ; 10% saccharose

Tampon de saccharose 30% : 0.1M NH<sub>4</sub>Cl ; 2.5mM MgAc ; 6mM  $\beta$ -M ; 20mM HEPES pH 7.5 ; 0.1mM benzamidine ; 0.5mM EDTA ; 30% saccharose

### 1/ Obtention des surnageants :

Reprendre les cellules dans 24mL de tampon A et les casser avec 24mL de billes en verre 2 fois 3 minutes avec un vibreur.

La solution est centrifugée 5 min à 4°C à 15300rpm, puis les surnageants sont rassemblés et les culots lavés avec 4mL de tampon A puis recentrifugés.

Les surnageants sont ultracentrifugés 75minutes à 4°C. Après avoir de nouveau récupéré le surnageant obtenu, celui-ci est réparti sur des coussins de saccharose, puis ultracentrifugé en 45Ti pendant 20h à 45000rpm à 4°C.

### 2/ Chromatographie sur colonne :

Le culot est repris dans du tampon A puis la solution est appliquée sur une colonne de chromatographie Séphacryl 400 préalablement équilibrée en tampon A. (0.3ml/min ; 3min/fractions).

Les fractions contenant les ribosomes sont rassemblées et l'éluat est dialysé 3h à 4°C contre le tampon de dialyse.

### 3/ Gradient de saccharose :

Le dialysat est appliqué sur des gradients de saccharose 10 et 30 % préalablement réalisés. Ces derniers sont ensuite ultracentrifugés 19h à 26100rpm à 4°C.

Les fractions contenant les sous-unités 30S et 50S sont récupérées grâce à une pompe péristaltique. Les fractions sont de nouveau ultracentrifugées 20h à 45000rpm à 4°C.

Le culot est ensuite repris avec du tampon A.

Un prélèvement de solution peut être réalisée à chaque étape afin de suivre la purification sur un gel natif 4% (1mL acrylamide 40% 19 :1 ; 2mL tampon TB 5X ; 50µL APS ; 10µL Temed ; QSP 10mL H<sub>2</sub>O)(tampon TB 5X : 30.29g TRIS ; 7.73g acide borique ; QSP 1L H<sub>2</sub>O).

## **6.3 Expression et purification des différents mutants d'aIF1**

Le gène codant pour aIF1 de *P. abyssi* est placé sous le contrôle d'un promoteur reconnu par l'ARN polymérase du phage T7 et du site opérateur de l'opéron lactose. Le gène codant l'ARN polymérase du phage T7 est présent sur l'ADN chromosomique des souches d'*E. coli* BL21-rosetta et est placé sous le contrôle du promoteur et du site opérateur de l'opéron lactose. Par ailleurs, les souches BL21-rosetta contiennent un plasmide dans lequel sont clonés LacI, gène codant pour le répresseur de l'opéron lactose, les gènes des ARNt rares d'*E. coli*, et un gène de résistance au chloramphénicol.

La souche BL21-rosetta d'*E. coli* électrocompétente est donc utilisée pour l'expression des différentes protéines. Après transformation de cette souche par l'un des plasmides d'intérêt, celle-ci est cultivée en milieu 2xTY (16g tryptone, 10g extrait de levure, 5g NaCl) complémenté de 50µg/mL d'ampicilline et 25µg/mL de chloramphénicol à 37°C sur la nuit. Lorsque les cellules sont en phase stationnaire, l'expression du gène d'intérêt est induite par l'ajout d'1mM d'IPTG puis les cellules sont laissées en culture 5h à 18°C.

Les milieux de culture sont centrifugés (5000g, 30min, 4°C), et séparés en deux de façon à obtenir deux aliquots équivalents à 500mL de culture. Les purifications des protéines surexprimées sont réalisées à partir d'un équivalent de 500mL de culture.

Chaque culot de culture est repris dans 20mL de tampon A10%B (tampon A : 10mM HEPES pH7.5, 1mM benzamidine, 1mM PMSF ; tampon B : 10mM HEPES pH7.5, 1mM benzamidine, 1mM PMSF, 1M NaCl). Les cellules sont désintégrées aux ultrasons, et l'extrait total est centrifugé 10 minutes à 20000rpm. La fraction de protéines solubles est chauffée 10 minutes à 70°C afin de faire précipiter les protéines bactériennes thermolabiles. Après centrifugation, le surnageant contenant les protéines restées solubles après chauffage (dont la protéine d'intérêt aIF1 mutant de *P. abyssi*) est appliqué sur une colonne échangeuse d'anion (Q-sépharose, GE Healthcare, 10mm×6cm) préalablement équilibrée en tampon A10%B. Les mutants aIF1\*, aIF1\*-YA33, aIF1\*-KA35, aIF1\*-RA32, aIF1\*-FA91, aIF1\*- ΔN, aIF1\*-polyAla ne s'accrochent pas sur cette résine. Par contre cette étape permet de débarrasser notre échantillon de la majorité des acides nucléiques qui interagissent fortement avec la résine. Les fractions correspondant au pic non retenu sur la colonne sont rassemblées et réappliquées sur une colonne échangeuse de cations (S-sépharose, GE Healthcare, 10mm×3cm) préalablement équilibrée en tampon A10%B.

Les protéines aIF1\*, aIF1\*-YA33 et aIF1\*- ΔN sont retenues sur cette colonne. Elles sont éluées par un gradient linéaire de NaCl de 100mM à 600mM (10 à 60%B) pendant 20 volumes de colonne. Les fractions correspondantes sont collectées, rassemblées, concentrées pour obtenir une concentration minimum de 100μM, puis stockées à -80°C.

Les protéines aIF1\*-RA32, aIF1\*-KA35 et aIF1\*-polyAla n'étant pas retenues sur la résine de la colonne S-sépharose, les fractions correspondant au pic non retenu de la colonne échangeuse d'anions sont rassemblées, concentrées dans 500μL, puis appliquées à 0,4mL/min sur un tamis moléculaire Superdex 75 (HR 10/30, GE Healthcare) équilibrée en tampon A10%B. Les fractions correspondants à aIF1 sont rassemblées, concentrées afin d'obtenir une concentration minimum de 100μM, puis stockées à -80°C.

La présence et l'évolution de la pureté des protéines au cours de la purification sont vérifiées par dépôt des différentes fractions sur gel de polyacrylamide 16% contenant du SDS.

	Poids moléculaire (Da)	Points isoélectrique	Elution sur S-sépharose	Passage sur SD75
<b>aIF1WT</b>	11469	6.61	20%B	non
<b>aIF1*</b>	11228	7.62	28%B	non
<b>aIF1*-RA32</b>	11343	5.74	10%B	oui
<b>aIF1*-YA33</b>	11336	6.35	20%B	non
<b>aIF1*-KA35</b>	11371	5.74	10%B	oui
<b>aIF1*-ΔN</b>	9119	7.62	20%B	non

<b>aIF1*-RA32-YA33</b>	1251	5.52	10%B	oui
<b>aIF1*-RA32-KA35</b>	11279	5.52	10%B	oui
<b>aIF1-RA32</b>	11384	5.92	10%B	oui
<b>aIF1-YA33</b>	11377	6.61	20%B	non
<b>aIF1-KA35</b>	11412	6.35	10%B	oui
<b>aIF1- ΔN</b>	9160	7.84	20%B	non
<b>aIF1-RA32-YA33</b>	11292	5.92	10%B	oui
<b>aIF1-RA32-KA35</b>	11320	5.92	10%B	oui

#### **6.4 Marquage des différents mutants d'aIF1 avec la coumarine**

Pour réaliser les marquages des variants d'aIF1, un dérivé maléimide de la coumarine est utilisé : le Mdcc ou 7-Diethylamino-3-(((2-Maleimidyl)ethyl)amino)carbonyl)coumarine. Le Mdcc (Fluka) est repris dans du diméthylformamide à une concentration de 20mM. La réaction de marquage des protéines est réalisée avec 5 ou 15 fois plus de Mdcc que de protéine à température ambiante, 25 ou 30°C pendant 2h sous agitation. Cette réaction correspond à la formation d'une liaison covalente entre le groupement maléimide du Mdcc et le groupement soufre du résidu cystéine d'aIF1. Par la suite, la fraction de protéine marquée est appliquée sur une colonne S-sépharose ou Superdex 75 selon le point isoélectrique du mutant aIF1. Cette purification permet d'obtenir des fractions contenant exclusivement la protéine d'intérêt marquée (mêlée à la fraction non marquée) et d'exclure le Mdcc resté libre.

Après concentration des fractions correspondantes, l'absorbance est mesurée à 430nm et à 280nm afin de calculer la concentration en protéines marquées ainsi que le rendement de marquage. En effet, le Mdcc absorbe les photons à 430nm avec un coefficient d'extinction molaire de  $46800\text{M}^{-1}\text{cm}^{-1}$  et à 280nm avec un coefficient de  $7470\text{M}^{-1}\text{cm}^{-1}$ . Par ailleurs, les protéines absorbent elles aussi les photons à 280nm. Pour connaître la concentration en protéine, il faut donc s'affranchir de la contribution du Mdcc à 280nm. La mesure de l'absorbance à 430nm permet de déterminer la concentration en fluorophore. A partir de cette donnée, on peut déterminer aussi la contribution du Mdcc à l'absorbance à 280nm. Finalement, la fraction d'aIF1 non marquée peut alors être déduite.

Absorbance (A) totale mesurée  $_{280\text{nm}} = A_{\text{aIF1}} + A_{\text{Mdcc}(280\text{nm})} = A_{\text{aIF1 marquée}}$

$A_{\text{aIF1 non marquée}} = A_{\text{totale}} - A_{\text{aIF1 marquée}}$

Rendement de marquage =  $[\text{aIF1}]_{\text{marquée}} \div [\text{aIF1}]_{\text{totale}} \times 100$

Exemple des calculs des concentrations en Mdcc et en protéine, ainsi que du rendement de marquage pour aIF1\* :

DO430nm mesurée = 2.25                      DO280nm mesurée = 0.57

Concentration en Mdcc =  $2,25 \div 46800 = 4,8 \cdot 10^{-5} \text{M} = 48 \mu\text{M}$

Contribution du Mdcc à la DO280nm =  $4,8 \cdot 10^{-5} \times 7470 = 0.359$

DO280nm aIF1\* =  $0,57 - 0,359 = 0,211$

Concentration d'aIF1\* =  $0,211 \div 2369 = 8,9 \cdot 10^{-5} \text{M} = 89 \mu\text{M}$

Rendement de marquage =  $(48 \div 89) \times 100 = 54\%$

Voir le détail de la réaction de marquage pour chaque mutant en annexe 2.

## 6.5 L'anisotropie de fluorescence

Pour chaque étude de mutant d'aIF1, le protocole reste le même, seules les concentrations en protéines pourront varier. Chaque titrage est effectué avec la petite sous-unité 30S du ribosome et un ARNm modèle comportant 26 bases de la région de démarrage de traduction du gène d'aEF1A (672nM de la sous-unité 30S de *P. abyssi*, 1340nM d'ARNm).

Les facteurs aIF1 WT ou mutants marqués préalablement avec la coumarine sont dilués dans un volume final de 220µL de tampon tween (200mM NaCl, 10 mM Hepès pH7.5, 10mM MgAc, 0.1% tween 20) afin d'obtenir une concentration minimum permettant une mesure de l'anisotropie de fluorescence. Le mélange transféré dans une cuve en quartz à 4 faces optiques Hellma 1cm×0.4cm est placé dans le spectrofluorimètre FP 8300 de JASCO. La cuve est chauffée à 50°C (compromis de température pour étudier la formation des complexes ribosomiaux de *P. abyssi* et limiter l'évaporation de l'échantillon) puis l'intensité de fluorescence est mesurée afin d'évaluer si le rapport signal/bruit est suffisamment élevé. La tension du photomultiplicateur (PM) est ajustée pour obtenir une intensité d'émission de fluorescence mesurée entre 1500 et 6000 unités arbitraires (la zone de linéarité se situant en théorie entre 1 et 10000) de manière à optimiser le rapport signal/bruit.

Les paramètres du spectrofluorimètre ont été ajustés de manière à maximiser la qualité des enregistrements :

- la longueur d'onde d'excitation est 428nm
- la bande passante d'excitation est de 2.5 ou 5nm

- l'émission de lumière est enregistrée à la longueur d'onde de 473nm (ce qui correspond au maximum d'émission de la coumarine)
- la bande passante d'émission est de à 10nm
- le facteur G (=IHV/IHH) est mesuré pour chaque mutant. Ce facteur permet de corriger la différence d'efficacité de transmission du photomultiplicateur pour la lumière polarisée verticale et horizontale

IHV = intensité de lumière émise avec une polarisation verticale pour une excitation polarisée verticalement

IHH = intensité de lumière émise une polarisation horizontale pour une excitation polarisée verticalement

L'anisotropie est définie par :  $A = (IVV - G \times IVH) / (IVV + 2G \times IVH)$

Chaque titrage d'aIF1 est réalisé avec 10 mesures d'anisotropie de fluorescence successives à t=1min et t=3min après chaque point de titrage. Lorsque l'anisotropie atteint une valeur plateau, un titrage retour est réalisé grâce à l'ajout de la protéine d'intérêt non marquée avec la coumarine à la concentration d'1,3μM.

Les résultats de mesure sont ensuite analysés afin d'obtenir la courbe de tendance de l'anisotropie en fonction de la concentration en ribosomes ajoutés. Les fluorescences d'aIF1 marqué libre ou complexé à la sous-unité 30S étant égales, l'anisotropie de fluorescence peut directement être reliée à la fraction d'aIF1 marqué liée au ribosome. En supposant un seul site de liaison d'aIF1 marqué au complexe 30S:ARNm (C), on a :

$$A_{\text{mesurée}} = A_{\text{aIF1marqué}} + [A_{\text{C:aIF1marqué}} - A_{\text{aIF1marqué}}] \left[ \frac{([aIF1] + [C]) - \sqrt{([aIF1] + [C])^2 - 4[aIF1][C]}}{2} \right]$$

Les valeurs de Kd,  $A_{\text{aIF1}}$  et  $A_{\text{C:aIF1}}$  en sont alors déduites, ainsi que les erreurs standard correspondantes, à l'aide du programme Origin.

Cette étude nécessite par ailleurs l'obtention préalable de facteurs de démarrage dans des fractions solubles pures, leur marquage avec de la coumarine, ainsi que l'obtention de la petite sous-unité 30S du ribosome de *P. abyssi* et d'un ARNm modèle.

## 6.6 L'expérience de toeprint

Le protocole est adaptable selon le nombre et la composition du jeu de facteurs de démarrage testés. L'ARNm modèle est obtenu par transcription *in vitro* à l'aide de l'ADN polymérase du phage T7.

La première étape consiste en l'hybridation de l'amorce sur l'ARNm modèle codant pour aEF1A de *P. abyssi* et ayant une région SD forte. Un mélange d'hybridation (« annealing ») est constitué de volume final  $2X \mu\text{L}$ ,  $X$  étant le nombre d'expériences à réaliser. Pour 15 expériences par exemple, il faut prévoir un volume de  $30\mu\text{L}$  de composition suivante : 1picomole d'ARNm purifié, 3pmoles d'oligonucléotide amorce,  $3\mu\text{L}$  de tampon 10X (600mM  $\text{NH}_4\text{Cl}$ , 100mM Hepes pH7.5, 70mM  $\beta$ -mercaptoéthanol),  $1\mu\text{L}$  de RNase out (40u/ $\mu\text{L}$ , Invitrogen) ; on complète à  $30\mu\text{L}$  par de l'eau MQ. La solution est chauffée 3min à  $60^\circ\text{C}$ , puis placée dans la glace avant l'ajout de  $1.8\mu\text{L}$  d'acétate de magnésium 1M pour atteindre la concentration de 60mM.

L'étape suivante est la constitution d'un prémix de composition suivante :  $1X \mu\text{L}$  de sous-unité 30S de *P. abyssi* diluées à 240nM dans son tampon de dilution (100mM  $\text{NH}_4\text{Cl}$ , 10.5mM  $\text{MgAc}$ , 6mM  $\beta$ -mercaptoéthanol, 0.5mM EDTA, 20mM Hepes pH7.5),  $2X \mu\text{L}$  d'annealing,  $0.25X \mu\text{L}$  de désoxyribonucléotides triphosphates à 3.75mM,  $0.25X \mu\text{L}$  GDP-NP à 5mM,  $0.25X \mu\text{L}$  de RNase out. Le volume final est égal à  $X \times 4.5\mu\text{L}$ . Le prémix est incubé 5min à  $51^\circ\text{C}$ . Les sous-unités 30S ont été purifiées au laboratoire par plusieurs étapes d'ultracentrifugation à partir de cellules de *P. abyssi*.

La dernière étape consiste à mélanger les différents facteurs. Le volume final est de  $10 \mu\text{L}$  par jeu de facteurs. Le prémix est présent à raison de  $4.5\mu\text{L}/\text{expériences}$ . S'ils sont présents, les facteurs aIF2, aIF1 et aIF1A de *P. abyssi* sont ajoutés à raison de  $1\mu\text{L}/\text{jeu}$  après avoir été dilués à  $2.5\mu\text{M}$  dans leur tampon de dilution (500mM  $\text{NaCl}$ , 10mM Hepes pH7.5, 10mM  $\beta$ -mercaptoéthanol). Le Met-ARNt (élongateur ou initiateur), s'il est présent, est ajouté à raison de  $1\mu\text{L}/\text{expériences}$  après avoir été dilué à  $10\mu\text{M}$  dans de l'eau MQ. Le tout est incubé une première fois 10min à  $51^\circ\text{C}$ . Après l'ajout de 0.25U de Reverse Transcriptase, les tests sont chauffés 15min à  $51^\circ\text{C}$ . La réaction est ensuite stoppée avec  $3\mu\text{L}$  de solution d'arrêt Licor (formamide + colorant).

Pour une base de 15 expériences, dans le volume réactionnel et avant l'ajout du tampon d'arrêt, les différents acteurs moléculaires du toeprint ont donc les concentrations suivantes : 250nM pour les facteurs de traduction, 200nM pour l'ARNt, 53.3nM pour les ribosomes et 1.5nM pour l'ARNm. Dans ces conditions, l'ARNm est saturé en facteurs, ARNt et ribosomes.

Les échantillons sont déposés sur gel de séquence. En parallèle, la réaction de séquençage de l'ADN d'aEF1A est déposée. L'image du gel de séquence obtenue grâce au logiciel Base ImageIR Data Collection version 4.0 est traitée à l'aide du logiciel Image J.

## 6.7 Isolation de petits ARNts

Préalablement, les cellules sont cassées, centrifugées, et le surnageant contenant les ARN test conservés pour la suite du protocole.

### materiels / tampons :

1. Oligonucléotide ADN complémentaire avec la totalité de l'ARNt cible. L'oligo est marqué en 3'-biotin et 5'-atto488. 2 à 3 adénosine sont rajoutées à l'extrémité 3' pour faciliter la liaison aux billes.

#### ARNt<sub>i</sub> *P. abyssi*:

Biotin-

AAATGGTAGCGGGGCGAGGATTTGAACCTCGGACCTCCGGTTATGAGCCCCGGCGGGCACTCCTAG  
CTGCCCCACCCCGCT-Atto488

#### ARNt<sub>i</sub> *S. solfataricus* :

Biotin-

AAATGGTAGCGGGGCGCTGGATTTGAACCAGGGACCTCGGGTTATGAGCCCCGCGCCTTACCAGG  
CTAGCCCACCCCGCT-Atto488

#### ARNt<sub>m</sub><sup>Met</sup> *P. abyssi*:

Biotin-

AAATGGTGCCGGGGCGGGGCTTCGAACCCCGGATCTCCCGTTATGAGCCGGGCGCTTTGACCAGG  
CTAAGTACCCCGGC-Atto488

#### ARNt<sub>m</sub><sup>Met</sup> *S. solfataricus* :

Biotin-

AAATGGTGCCGCCGCGGGGATTTGAACCCCGGACAACCGGATTATGAGTCCGGCGCTTAACCAGGC  
TGAGTACGGCGGC-Atto488

2. MyOneT1 Dynabeads (des billes magnétiques recouvertes de streptavidine)
3. Tampon B&W (1X)

Tampon 2X B&W : 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl

4. SSC buffer : 20X, 5X, 1X, 0.1X

20X : 3 M NaCl, 300 mM trisodiumcitrate (pH 7.0)

5. MilliQ H<sub>2</sub>O

### 1 : Hybridation

Les calculs sont faits pour 25  $\mu\text{L}$  de MyOneT1 Dynabeads (capacité de liaison de 100 pmol maximum pour des oligonucléotides biotinylés)

	<b>Stock</b>	<b>concentration finale</b>	<b>Volume (<math>\mu\text{L}</math>)</b>
<b>Oligo-ADN (biotinylé)</b>	100 $\mu\text{M}$	1 $\mu\text{M}$	1 $\mu\text{L}$
<b>RNA/tRNA Total</b>	*	$\sim 100$ pmol	*
<b>Tampon SSC</b>	20X	5X	25 $\mu\text{L}$
<b>MilliQ H<sub>2</sub>O</b>			QSP 100 $\mu\text{L}$

Denaturation : 3 min à 90 °C

Hybridation : 10 min à 65 °C

Conserver à température ambiante, protégé de la lumière jusqu'à ce que les billes soient prêtes à être utilisées (10-15min)

## 2 : Préparation des billes

Resuspendre les billes par pipetage et vortexage à demi-vitesse.

Pour séparer les billes du supernageant : centrifuger doucement, placer sur une grille magnétique  $\sim 2$  min, retirez soigneusement surnageant (les tubes restent encore sur la grille magnétique), essayez de ne pas toucher le culot avec la pointe de la pipette. Resuspendre les billes.

Préparation des billes:

Transférer les billes dans un nouvel eppendorf. Ajouter un volume égal de tampon 1X B & W, resuspendre et retirer le surnageant. Répétez 3 fois. Ajouter un volume égal de tampon 5X SSC, resuspendre, retirer le surnageant (ou préparer des aliquots et retirer le surnageant après).

Immédiatement passer à l'étape 3

## 3 : Immobilisation de la cible ARN-AND hybride

Ajouter les échantillons hybrides (100 $\mu\text{L}$  à température ambiante) aux bille, vortexer doucement et laisser agiter 30min à 25°C (600rpm)

## 4 : Elimination des ARNt non-cible restant

Centrifuge doucement les billes, placer les tubes sur le rack magnétique pendant 2min et retirer le supernageant (100 $\mu\text{L}$ , contenant les ARNt non cibles). Laver les billes dans 50 $\mu\text{L}$  de tampon 1X SSC, puis 3 fois dans 25 $\mu\text{L}$  de 0.1X SSC. Retirer le supernageant et passer immédiatement à l'étape 5.

## 5 : Elution de l'ARNt cible

Les billes sont resuspendues dans 25µL d'eau milliQ. Chauffer les billes à 75°C 3min, centrifuger et placer le tube sur le rack magnétique. Enlever le supernageant (contenant l'ARNt cible) et transférer dans un nouvel eppendorf (il peut être ensuite conserve à -20°C)

#### 6 : Digestion des oligo-ADN restant

Des oligo-AND peuvent être toujours presents malgré l'étape d'éluion et pourrait interferer avec les prochaines experiences. Pour les retirer entièrement, nous ajoutons de la DNase 1.

	<b>stock</b>	<b>volume</b>
<b>RNA/DNA mix élué</b>		25 µL
<b>Tampon DNase 1 (+MgCl<sub>2</sub>)</b>	10X	3 µL
<b>DNase 1</b>	50 U/µL	2 µL

Incubation : 2 h à 37 °C (il peut être ensuite conserve à -20°C)

#### 7 : Analyse PAGE et elution de la cible ARNt

Faire migrer sur un gel PAGE denature 10% un extrait de l'éluat digéré à la DNase I, les oligos et l'extrait total.

Electrophorèse : 16 W, ca. 1.5 h

(oligo : 1 µL à 10 µM ; extrait total : 1 µg)

Scan 1 :

gel non coloré : seul les oligos sont détectables (contrôle-scan pour la digenstion par la DNase1)

(Excitation: 532 nm / émission: 526 SP (<526 nm))

Coloration: GelRed (ou SYBR Gold)

Scan 2: gel coloré : excision des bandes d'ARNt cible

[Excitation: 532 nm ou 610 nm / émission: 526 SP (<526 nm)]

bande excisée: écrasez bien et éluer en volume approprié (~ 400 µL) de 0,5 M d'acétate d'ammonium pendant une nuit (20 ° C, 550 rpm). Centrifuger les morceau de gel, filtrer le surnageant (Nanosep MF 0,45 µm). Réaliser une précipitation éthanol des ARNt (ajouter 2,5 fois le volume d'éthanol pur (-80 ° C)). Dissoudre dans un volume approprié d'eau MilliQ et conserver à -20 ° C

# 7 Bibliographie

- Algire, M. A., D. Maag, et J. R. Lorsch. 2005. « Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation ». *Mol Cell* 20 (2): 251- 62.
- Allen, G. S., A. Zavialov, R. Gursky, M. Ehrenberg, et J. Frank. 2005. « The Cryo-EM Structure of a Translation Initiation Complex from Escherichia Coli ». *Cell* 121 (5): 703- 12. doi:10.1016/j.cell.2005.03.023.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, et D. J. Lipman. 1990. « Basic Local Alignment Search Tool ». *Journal of Molecular Biology* 215 (3): 403- 10. doi:10.1016/S0022-2836(05)80360-2.
- Anderson, J., L. Phan, R. Cuesta, B. A. Carlson, M. Pak, K. Asano, G. R. Björk, M. Tamame, et A. G. Hinnebusch. 1998. « The Essential Gcd10p-Gcd14p Nuclear Complex Is Required for 1-Methyladenosine Modification and Maturation of Initiator Methionyl-tRNA ». *Genes & Development* 12 (23): 3650- 62.
- Antoun, Ayman, Michael Y. Pavlov, Kerstin Andersson, Tanel Tenson, et Måns Ehrenberg. 2003. « The Roles of Initiation Factor 2 and Guanosine Triphosphate in Initiation of Protein Synthesis ». *The EMBO Journal* 22 (20): 5593- 5601. doi:10.1093/emboj/cdg525.
- Armache, J. P., A. M. Anger, V. Marquez, S. Franckenberg, T. Frohlich, E. Villa, O. Berninghausen, et al. 2013. « Promiscuous Behaviour of Archaeal Ribosomal Proteins: Implications for Eukaryotic Ribosome Evolution ». *Nucleic Acids Res.* 41 (2): 1284- 93. doi:10.1093/nar/gks1259.
- Ban, N., P. Nissen, J. Hansen, P. B. Moore, et T. A. Steitz. 2000. « The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution ». *Science (New York, N.Y.)* 289 (5481): 905- 20.
- Ban, Nenad, Roland Beckmann, Jamie HD Cate, Jonathan D Dinman, François Dragon, Steven R Ellis, Denis LJ Lafontaine, et al. 2014. « A new system for naming ribosomal proteins ». *Current opinion in structural biology* 24 (février): 165- 69. doi:10.1016/j.sbi.2014.01.002.
- Barraud, P., E. Schmitt, Y. Mechulam, F. Dardel, et C. Tisne. 2008. « A Unique Conformation of the Anticodon Stem-Loop Is Associated with the Capacity of tRNA<sup>Met</sup> to Initiate Protein Synthesis ». *Nucleic Acids Res* 36 (15): 4894- 4901. doi:10.1093/nar/gkn462.
- Bartig, D., K. Lemkemeier, J. Frank, F. Lottspeich, et F. Klink. 1992. « The Archaeobacterial Hypusine-Containing Protein. Structural Features Suggest Common Ancestry with Eukaryotic Translation Initiation Factor 5A ». *European Journal of Biochemistry / FEBS* 204 (2): 751- 58.
- Battiste, J. L., T. V. Pestova, C. U. Hellen, et G. Wagner. 2000. « The eIF1A Solution Structure Reveals a Large RNA-Binding Surface Important for Scanning Function ». *Mol Cell* 5 (1): 109- 19.
- Battle, D. J., et J. A. Doudna. 2002. « Specificity of RNA-RNA helix recognition ». *Proc Natl Acad Sci U S A* 99 (18): 11676- 81. doi:10.1073/pnas.182221799.
- Benelli, Dario, Enzo Maone, et Paola Londei. 2003. « Two Different Mechanisms for ribosome/mRNA Interaction in Archaeal Translation Initiation ». *Molecular Microbiology* 50 (2): 635- 43.
- Ben-Shem, A., N. Garreau de Loubresse, S. Melnikov, L. Jenner, G. Yusupova, et M. Yusupov. 2011. « The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution ». *Science* 334 (6062): 1524- 29. doi:10.1126/science.1212642.
- Berk, Veysel, Wen Zhang, Raj D. Pai, Jamie H. D. Cate, et Jamie H. Doudna. 2006. « Structural Basis for mRNA and tRNA Positioning on the Ribosome ». *Proceedings of the National Academy of Sciences of the United States of America* 103 (43): 15830- 34. doi:10.1073/pnas.0607541103.
- Biou, V., F. Shu, et V. Ramakrishnan. 1995. « X-Ray Crystallography Shows That Translational Initiation Factor IF3 Consists of Two Compact Alpha/beta Domains Linked by an Alpha-Helix ». *The EMBO Journal* 14 (16): 4056- 64.
- Boileau, G., P. Butler, J. W. Hershey, et R. R. Traut. 1983. « Direct Cross-Links between Initiation Factors 1, 2, and 3 and Ribosomal Proteins Promoted by 2-Iminoethanol ». *Biochemistry* 22 (13): 3162- 70.

- Borck, G., B. S. Shin, B. Stiller, A. Mimouni-Bloch, H. Thiele, J. R. Kim, M. Thakur, et al. 2012. « eIF2gamma mutation that disrupts eIF2 complex integrity links intellectual disability to impaired translation initiation ». *Mol Cell* 48 (4): 641- 46. doi:10.1016/j.molcel.2012.09.005.
- Britten, R. J., et R. B. Roberts. 1960. « High-Resolution Density Gradient Sedimentation Analysis ». *Science (New York, N.Y.)* 131 (3392): 32- 33. doi:10.1126/science.131.3392.32.
- Brown, J. W., C. J. Daniels, et J. N. Reeve. 1989. « Gene Structure, Organization, and Expression in Archaeobacteria ». *Critical Reviews in Microbiology* 16 (4): 287- 338. doi:10.3109/10408418909105479.
- Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, et al. 1996. « Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii* ». *Science* 273 (5278): 1058- 73.
- Bycroft, M., T. J. Hubbard, M. Proctor, S. M. Freund, et A. G. Murzin. 1997. « The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold ». *Cell* 88 (2): 235- 42.
- Carter, A. P., W. M. Clemons Jr., D. E. Brodersen, R. J. Morgan-Warren, T. Hartsch, B. T. Wimberly, et V. Ramakrishnan. 2001. « Crystal structure of an initiation factor bound to the 30S ribosomal subunit ». *Science* 291 (5503): 498- 501.
- Castilho-Valavicius, B., H. Yoon, et T. F. Donahue. 1990. « Genetic characterization of the *Saccharomyces cerevisiae* translational initiation suppressors *sui1*, *sui2* and *SUI3* and their effects on *HIS4* expression ». *Genetics* 124 (3): 483- 95.
- Cheung, Y. N., D. Maag, S. F. Mitchell, C. A. Fekete, M. A. Algire, J. E. Takacs, N. Shirokikh, T. Pestova, J. R. Lorsch, et A. G. Hinnebusch. 2007. « Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo ». *Genes Dev* 21 (10): 1217- 30.
- Cigan, A. M., E. K. Pabich, L. Feng, et T. F. Donahue. 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. USA* 86 (8): 2784- 88.
- Clemens, M. J., et U. A. Bommer. 1999. « Translational Control: The Cancer Connection ». *The International Journal of Biochemistry & Cell Biology* 31 (1): 1- 23.
- Clemens, M.J. 1990. « Does protein phosphorylation play a role in translational control by eukaryotic aminoacyl-tRNA synthetases? » *Trends in Biochemical Sciences* 15: 172- 75.
- Costa-Mattioli, Mauro, Delphine Gobert, Elad Stern, Karine Gamache, Rodney Colina, Claudio Cuello, Wayne Sossin, et al. 2007. « eIF2alpha Phosphorylation Bidirectionally Regulates the Switch from Short- to Long-Term Synaptic Plasticity and Memory ». *Cell* 129 (1): 195- 206. doi:10.1016/j.cell.2007.01.050.
- Culver, G. M., G. M. Heilek, et H. F. Noller. 1999. « Probing the rRNA Environment of Ribosomal Protein S5 across the Subunit Interface and inside the 30 S Subunit Using Tethered Fe(II) ». *Journal of Molecular Biology* 286 (2): 355- 64. doi:10.1006/jmbi.1998.2483.
- Culver, G. M., et H. F. Noller. 1999. « Efficient Reconstitution of Functional *Escherichia Coli* 30S Ribosomal Subunits from a Complete Set of Recombinant Small Subunit Ribosomal Proteins ». *RNA (New York, N.Y.)* 5 (6): 832- 43.
- Cummings, H. S., et J. W. Hershey. 1994. « Translation Initiation Factor IF1 Is Essential for Cell Viability in *Escherichia Coli* ». *Journal of Bacteriology* 176 (1): 198- 205.
- Dahlquist, K. D., et J. D. Puglisi. 2000. « Interaction of Translation Initiation Factor IF1 with the *E. Coli* Ribosomal A Site ». *Journal of Molecular Biology* 299 (1): 1- 15. doi:10.1006/jmbi.2000.3672.
- Dallas, A., et H. F. Noller. 2001. « Interaction of Translation Initiation Factor 3 with the 30S Ribosomal Subunit ». *Molecular Cell* 8 (4): 855- 64.
- Datsenko, K. A., et B. L. Wanner. 2000. « One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products ». *Proc Natl Acad Sci U S A* 97 (12): 6640- 45. doi:10.1073/pnas.120163297.
- De Bellis, D., D. Liveris, D. Goss, S. Ringquist, et I. Schwartz. 1992. « Structure-Function Analysis of *Escherichia Coli* Translation Initiation Factor IF3: Tyrosine 107 and Lysine 110 Are Required for Ribosome Binding ». *Biochemistry* 31 (48): 11984- 90.

- de Cock, E., M. Springer, et F. Dardel. 1999. « The Interdomain Linker of Escherichia Coli Initiation Factor IF3: A Possible Trigger of Translation Initiation Specificity ». *Molecular Microbiology* 32 (1): 193- 202.
- Dettman, G. L., et W. M. Stanley. 1972. « Recognition of Eukaryotic Initiator tRNA by an Initiation Factor and the Transfer of the Methionine Moiety into Peptide Linkage ». *Biochimica Et Biophysica Acta* 287 (1): 124- 33.
- Dever, T. E., C. L. Wei, L. A. Benkowski, K. Browning, W. C. Merrick, et J. W. Hershey. 1994. « Determination of the Amino Acid Sequence of Rabbit, Human, and Wheat Germ Protein Synthesis Factor eIF-4C by Cloning and Chemical Sequencing ». *The Journal of Biological Chemistry* 269 (5): 3212- 18.
- Donahue, T. F., A. M. Cigan, E. K. Pabich, et B. C. Valavicius. 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 (5): 621- 32.
- Dorris, D. R., F. L. Erickson, et E. M. Hannig. 1995. « Mutations in GCD11, the structural gene for eIF-2 gamma in yeast, alter translational regulation of GCN4 and the selection of the start site for protein synthesis ». *EMBO J.* 14 (10): 2239- 49.
- Erickson, F. L., et E. M. Hannig. 1996. « Ligand interactions with eukaryotic translation initiation factor 2: role of the gamma-subunit ». *EMBO J...* 15 (22): 6311- 6320.
- Fekete, C. A., D. J. Applefield, S. A. Blakely, N. Shirokikh, T. Pestova, J. R. Lorsch, et A. G. Hinnebusch. 2005. « The eIF1A C-Terminal Domain Promotes Initiation Complex Assembly, Scanning and AUG Selection in Vivo ». *EMBO J* 24 (20): 3588- 3601. doi:10.1038/sj.emboj.7600821.
- Fletcher, C. M., T. V. Pestova, C. U. Hellen, et G. Wagner. 1999. « Structure and interactions of the translation initiation factor eIF1 ». *Embo J* 18 (9): 2631- 37.
- Garcia, C., P. L. Fortier, S. Blanquet, J. Y. Lallemand, et F. Dardel. 1995a. « 1H and 15N Resonance Assignments and Structure of the N-Terminal Domain of Escherichia Coli Initiation Factor 3 ». *European Journal of Biochemistry / FEBS* 228 (2): 395- 402.
- Garcia, C., P.L. Fortier, S. Blanquet, J.Y. Lallemand, et F. Dardel. 1995b. « 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- 59.
- Greber, Basil J., Philipp Bieri, Marc Leibundgut, Alexander Leitner, Ruedi Aebersold, Daniel Boehringer, et Nenad Ban. 2015. « Ribosome. The Complete Structure of the 55S Mammalian Mitochondrial Ribosome ». *Science (New York, N.Y.)* 348 (6232): 303- 8. doi:10.1126/science.aaa3872.
- Green, R., et H. F. Noller. 1999. « Reconstitution of Functional 50S Ribosomes from in Vitro Transcripts of Bacillus Stearothermophilus 23S rRNA ». *Biochemistry* 38 (6): 1772- 79. doi:10.1021/bi982246a.
- Gualerzi, C. O., M. Severini, R. Spurio, A. La Teana, et C. L. Pon. 1991. « Molecular dissection of translation initiation factor IF2. Evidence for two structural and functional domains ». *J. Biol. Chem.* 266 (25): 16356- 62.
- Gualerzi, C. O., R. Spurio, A. La Teana, R. Calogero, B. Celano, et C. L. Pon. 1989. « Site-Directed Mutagenesis of Escherichia Coli Translation Initiation Factor IF1. Identification of the Amino Acid Involved in Its Ribosomal Binding and Recycling ». *Protein Engineering* 3 (2): 133- 38.
- Gualerzi, C., G. Risuleo, et C. L. Pon. 1977. « Initial Rate Kinetic Analysis of the Mechanism of Initiation Complex Formation and the Role of Initiation Factor IF-3 ». *Biochemistry* 16 (8): 1684- 89.
- Guillon, J. M., S. Heiss, J. Soutourina, Y. Mechulam, S. Laalami, M. Grunberg-Manago, et S. Blanquet. 1996. « Interplay of Methionine tRNAs with Translation Elongation Factor Tu and Translation Initiation Factor 2 in Escherichia Coli ». *J Biol Chem* 271 (37): 22321- 25.
- Guillon, J.M., Y. Mechulam, S. Blanquet, et G. Fayat. 1993. « Importance of formylability and anticodon stem sequence to give tRNAMet an initiator identity in Escherichia coli ». *J. Bacteriol.* 175: 4507- 14.

- Guillon, J.M., Y. Mechulam, J.M. Schmitter, S. Blanquet, et G. Fayat. 1992. « Disruption of the gene for Met-tRNA<sup>fMet</sup> formyltransferase severely impairs growth of *Escherichia coli* ». *J. Bacteriol.* 174: 4294- 4301.
- Harms, J., F. Schluenzen, R. Zarivach, A. Bashan, S. Gat, I. Agmon, H. Bartels, F. Franceschi, et A. Yonath. 2001. « High Resolution Structure of the Large Ribosomal Subunit from a Mesophilic Eubacterium ». *Cell* 107 (5): 679- 88. doi:S0092-8674(01)00546-3 [pii].
- Hartz, D., J. Binkley, T. Hollingsworth, et L. Gold. 1990. « Domains of Initiator tRNA and Initiation Codon Crucial for Initiator tRNA Selection by *Escherichia Coli* IF3 ». *Genes & Development* 4 (10): 1790- 1800.
- Hasenöhrl, D., D. Benelli, A. Barbazza, P. Londei, et U. Bläsi. 2006. « *Sulfolobus Solfataricus* Translation Initiation Factor 1 Stimulates Translation Initiation Complex Formation ». *RNA* 12 (4): 674- 82. doi:10.1261/rna.2289306.
- Hasenöhrl, D., A. Fabbretti, P. Londei, C. O. Gualerzi, et U. Bläsi. 2009. « Translation Initiation Complex Formation in the Crenarchaeon *Sulfolobus Solfataricus* ». *RNA* 15 (12): 2288- 98. doi:10.1261/rna.1662609.
- Hinnebusch, A. G. 1993. « Gene-specific translational control of the yeast GCN4 gene by phosphorylation of eukaryotic initiation factor 2 ». *Mol Microbiol* 10 (2): 215- 23.
- Hinnebusch, A. G., et J. R. Lorsch. 2012. « The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges ». *Cold Spring Harb. Perspect. Biol.* 4 (10): 1- 25. doi:10.1101/cshperspect.a011544.
- Holmberg, L., et H. F. Noller. 1999. « Mapping the Ribosomal RNA Neighborhood of Protein L11 by Directed Hydroxyl Radical Probing ». *Journal of Molecular Biology* 289 (2): 223- 33. doi:10.1006/jmbi.1999.2706.
- Hori, Hiroyuki. 2014. « Methylated Nucleosides in tRNA and tRNA Methyltransferases ». *Frontiers in Genetics* 5: 144. doi:10.3389/fgene.2014.00144.
- Huang, H. K., H. Yoon, E. M. Hannig, et T. F. Donahue. 1997. « GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae* ». *Genes Dev.* 11 (18): 2396- 2413.
- Jovine, L., S. Djordjevic, et D. Rhodes. 2000. « The crystal structure of yeast phenylalanine tRNA at 2.0 Å resolution: cleavage by Mg(2+) in 15-year old crystals ». *J Mol Biol* 301 (2): 401- 14. doi:10.1006/jmbi.2000.3950.
- Kapp, L. D., S. E. Koltz, et J. R. Lorsch. 2006. « Yeast Initiator tRNA Identity Elements Cooperate to Influence Multiple Steps of Translation Initiation ». *RNA* 12 (5): 751- 64. doi:10.1261/rna.2263906.
- Kasperaitis, M. A., H. O. Voorma, et A. A. Thomas. 1995. « The Amino Acid Sequence of Eukaryotic Translation Initiation Factor 1 and Its Similarity to Yeast Initiation Factor SUI1 ». *FEBS Letters* 365 (1): 47- 50.
- Khatter, Heena, Alexander G. Myasnikov, S. Kundhavai Natchiar, et Bruno P. Klaholz. 2015. « Structure of the Human 80S Ribosome ». *Nature* 520 (7549): 640- 45. doi:10.1038/nature14427.
- Kim, S. H., G. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, et Alexander Rich. 1973. « Unit cell transformations in yeast phenylalanine transfer RNA crystals ». *Journal of Molecular Biology* 75 (2): 429- 32. doi:10.1016/0022-2836(73)90032-6.
- Kozak, M. 1981. « Possible Role of Flanking Nucleotides in Recognition of the AUG Initiator Codon by Eukaryotic Ribosomes ». *Nucleic Acids Research* 9 (20): 5233- 52.
- . 1986. « Point Mutations Define a Sequence Flanking the AUG Initiator Codon That Modulates Translation by Eukaryotic Ribosomes ». *Cell* 44 (2): 283- 92.
- Kozak, Marilyn. 2002. « Pushing the Limits of the Scanning Mechanism for Initiation of Translation ». *Gene* 299 (1- 2): 1- 34.
- Kycia, J. H., V. Biou, F. Shu, S. E. Gerchman, V. Graziano, et V. Ramakrishnan. 1995. « Prokaryotic Translation Initiation Factor IF3 Is an Elongated Protein Consisting of Two Crystallizable Domains ». *Biochemistry* 34 (18): 6183- 87.

- Kyrpides, N. C., et C. R. Woese. 1998. « Universally conserved translation initiation factors ». *Proc. Natl. Acad. Sci. USA* 95 (1): 224- 8.
- Lee, C.P., B.L. Seong, et U.L. RajBhandary. 1991. « Structural and sequence elements important for recognition of Escherichia coli formylmethionine tRNA by methionyl-tRNA transformylase are clustered in the acceptor stem ». *J. Biol. Chem.* 266 (27): 18012- 17.
- Levin, D. H., D. Kyner, et G. Acs. 1971. « Messenger Activity in Mammalian Cell-Free Extracts of Reovirus Single-Stranded RNA Prepared in Vitro ». *Biochemical and Biophysical Research Communications* 42 (3): 454- 61.
- . 1973. « Protein Initiation in Eukaryotes: Formation and Function of a Ternary Complex Composed of a Partially Purified Ribosomal Factor, Methionyl Transfer RNA, and Guanosine Triphosphate ». *Proc. Natl. Acad. Sci. USA* 70 (1): 41- 45.
- Li, S., N.V. Kumar, U. Varsney, et U.L. RajBhandary. 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- 28.
- Lomakin, I. B., V. G. Kolupaeva, A. Marintchev, G. Wagner, et T. V. Pestova. 2003. « Position of Eukaryotic Initiation Factor eIF1 on the 40S Ribosomal Subunit Determined by Directed Hydroxyl Radical Probing ». *Genes Dev.* 17 (22): 2786- 97. doi:10.1101/gad.1141803.
- Lomakin, I. B., et T. A. Steitz. 2013. « The Initiation of Mammalian Protein Synthesis and mRNA Scanning Mechanism ». *Nature* 500 (juillet): 307- 11. doi:10.1038/nature12355.
- Lorsch, J., et T. Dever. 2010. « Molecular View of 43S Complex Formation and Start Site Selection in Eukaryotic Translation Initiation ». *J. Biol. Chem.* 285 (mai): 21203- 7. doi:10.1074/jbc.R110.119743.
- Maag, D., M. A. Algire, et J. R. Lorsch. 2006. « Communication between eukaryotic translation initiation factors 5 and 1A within the ribosomal pre-initiation complex plays a role in start site selection ». *J. Mol. Biol.* 356 (3): 724- 37.
- Maag, D., C. A. Fekete, Z. Gryczynski, et J. R. Lorsch. 2005. « A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon ». *Mol. Cell* 17 (2): 265- 75.
- Maag, D., et J. R. Lorsch. 2003. « Communication between eukaryotic translation initiation factors 1 and 1A on the yeast small ribosomal subunit ». *J. Mol. Biol.* 330 (5): 917- 24.
- Mandal, N., et U.L. Rajbhandary. 1992. « Escherichia coli B lacks one of the 2 initiator transfer RNA species present in E. coli K-12 ». *Journal of Bacteriology* 174 (23): 7827- 30.
- Marck, C., et H. Grosjean. 2002. « tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features ». *RNA* 8 (10): 1189- 1232.
- Marintchev, Assen, Victoria G. Kolupaeva, Tatyana V. Pestova, et Gerhard Wagner. 2003. « Mapping the Binding Interface between Human Eukaryotic Initiation Factors 1A and 5B: A New Interaction between Old Partners ». *Proceedings of the National Academy of Sciences of the United States of America* 100 (4): 1535- 40. doi:10.1073/pnas.0437845100.
- Martin-Marcos, P., J. Nanda, R. E. Luna, G. Wagner, J. R. Lorsch, et A. G. Hinnebusch. 2013. « Beta-Hairpin Loop of eIF1 Mediates 40S Ribosome Binding to Regulate Initiator tRNAMet Recruitment and Accuracy of AUG Selection in Vivo ». *J. Biol. Chem.*, juillet, 27546- 62. doi:10.1074/jbc.M113.498642.
- Mayer, Christine, Caroline Köhrer, Eimear Kenny, Carsten Prusko, et Uttam L. RajBhandary. 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 (17): 4787- 99. doi:10.1021/bi034011r.
- McCutcheon, J. P., R. K. Agrawal, S. M. Philips, R. A. Grassucci, S. E. Gerchman, W. M. Clemons, V. Ramakrishnan, et J. Frank. 1999. « Location of Translational Initiation Factor IF3 on the Small Ribosomal Subunit ». *Proceedings of the National Academy of Sciences of the United States of America* 96 (8): 4301- 6.

- Mechulam, Yves, Sylvain Blanquet, et Emmanuelle Schmitt. 2011. « Translation Initiation ». *EcoSal Plus* 4 (2). doi:10.1128/ecosalplus.4.2.2.
- Milón, Pohl, Cristina Maracci, Liudmila Filonava, Claudio O. Gualerzi, et Marina V. Rodnina. 2012. « Real-Time Assembly Landscape of Bacterial 30S Translation Initiation Complex ». *Nature Structural & Molecular Biology* 19 (6): 609- 15. doi:10.1038/nsmb.2285.
- Mitchell, Sarah F., et Jon R. Lorsch. 2008. « Should I Stay or Should I Go? Eukaryotic Translation Initiation Factors 1 and 1A Control Start Codon Recognition ». *The Journal of Biological Chemistry* 283 (41): 27345- 49. doi:10.1074/jbc.R800031200.
- Moazed, D., R. R. Samaha, C. Gualerzi, et H. F. Noller. 1995. « Specific Protection of 16 S rRNA by Translational Initiation Factors ». *Journal of Molecular Biology* 248 (2): 207- 10.
- Moreau, M., E. de Cock, P. L. Fortier, C. Garcia, C. Albaret, S. Blanquet, J. Y. Lallemand, et F. Dardel. 1997. « Heteronuclear NMR Studies of E. Coli Translation Initiation Factor IF3. Evidence That the Inter-Domain Region Is Disordered in Solution ». *Journal of Molecular Biology* 266 (1): 15- 22.
- Moreno, J. M., L. Drskjotersen, J. E. Kristensen, K. K. Mortensen, et H. U. Sperling-Petersen. 1999. « Characterization of the Domains of E. Coli Initiation Factor IF2 Responsible for Recognition of the Ribosome ». *FEBS Letters* 455 (1- 2): 130- 34.
- Nissen, P., J. A. Ippolito, N. Ban, P. B. Moore, et T. A. Steitz. 2001. « RNA tertiary interactions in the large ribosomal subunit: the A-minor motif ». *Proc Natl Acad Sci U S A* 98 (9): 4899- 4903. doi:10.1073/pnas.081082398.
- Olsson, C. L., M. Graffe, M. Springer, et J. W. Hershey. 1996. « Physiological Effects of Translation Initiation Factor IF3 and Ribosomal Protein L20 Limitation in Escherichia Coli ». *Molecular & General Genetics: MGG* 250 (6): 705- 14.
- Panvert, Michel, Etienne Dubiez, Lea Arnold, Javier Perez, Yves Mechulam, Wolfgang Seufert, et Emmanuelle Schmitt. 2015. « Cdc123, a Cell Cycle Regulator Needed for eIF2 Assembly, Is an ATP-Grasp Protein with Unique Features ». *Structure (London, England: 1993)* 23 (9): 1596- 1608. doi:10.1016/j.str.2015.06.014.
- Passmore, L. A., T. M. Schmeing, D. Maag, D. J. Applefield, M. G. Acker, M. A. Algire, J. R. Lorsch, et V. Ramakrishnan. 2007. « The Eukaryotic Translation Initiation Factors eIF1 and eIF1A Induce an Open Conformation of the 40S Ribosome ». *Mol. Cell* 26 (1): 41- 50.
- Pestova, T. V., S. I. Borukhov, et C. U. Hellen. 1998. « Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons ». *Nature* 394 (6696): 854- 9.
- Pestova, T. V., et V. G. Kolupaeva. 2002. « The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection ». *Genes Dev* 16 (22): 2906- 22.
- Petrelli, Dezemona, Anna LaTeana, Cristiana Garofalo, Roberto Spurio, Cynthia L. Pon, et Claudio O. Gualerzi. 2001. « Translation initiation factor IF3: two domains, five functions, one mechanism? ». *The EMBO Journal* 20 (16): 4560- 69. doi:10.1093/emboj/20.16.4560.
- Pioletti, M., F. Schlünzen, J. Harms, R. Zarivach, M. Glühmann, H. Avila, A. Bashan, et al. 2001. « Crystal Structures of Complexes of the Small Ribosomal Subunit with Tetracycline, Edeine and IF3 ». *The EMBO Journal* 20 (8): 1829- 39. doi:10.1093/emboj/20.8.1829.
- Pon, C. L., M. Paci, R. T. Pawlik, et C. O. Gualerzi. 1985. « Structure-Function Relationship in Escherichia Coli Initiation Factors. Biochemical and Biophysical Characterization of the Interaction between IF-2 and Guanosine Nucleotides ». *The Journal of Biological Chemistry* 260 (15): 8918- 24.
- Rabl, J., M. Leibundgut, S. F. Ataide, A. Haag, et N. Ban. 2011. « Crystal Structure of the Eukaryotic 40S Ribosomal Subunit in Complex with Initiation Factor 1 ». *Science* 331 (6018): 730- 36. doi:10.1126/science.1198308.
- Ramakrishnan, V. 2002. « Ribosome Structure and the Mechanism of Translation ». *Cell* 108 (4): 557- 72.
- Robertus, J. D., J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. Clark, et A. Klug. 1974. « Structure of Yeast Phenylalanine tRNA at 3 Å Resolution ». *Nature* 250 (467): 546- 51.

- Sacerdot, C., G. Fayat, P. Dessen, M. Springer, J. A. Plumbridge, M. Grunberg-Manago, et S. Blanquet. 1982. « Sequence of a 1.26-Kb DNA Fragment Containing the Structural Gene for E.coli Initiation Factor IF3: Presence of an AUU Initiator Codon ». *The EMBO Journal* 1 (3): 311- 15.
- Saini, A. K., J. S. Nanda, J. R. Lorsch, et A. G. Hinnebusch. 2009. « Regulatory Elements in eIF1A Control the Fidelity of Start Codon Selection by Modulating tRNA(i)(Met) Binding to the Ribosome ». *Genes Dev.* 24 (1): 97- 110. doi:10.1101/gad.1871910.
- Sands, J. F., H. S. Cummings, C. Sacerdot, L. Dondon, M. Grunberg-Manago, et J. W. Hershey. 1987. « Cloning and Mapping of infA, the Gene for Protein Synthesis Initiation Factor IF1 ». *Nucleic Acids Research* 15 (13): 5157- 68.
- Schmitt, E., M. Naveau, et Y. Mechulam. 2010. « Eukaryotic and Archaeal Translation Initiation Factor 2: A Heterotrimeric tRNA Carrier ». *FEBS Lett.* 584 (2): 405- 12. doi:10.1016/j.febslet.2009.11.002.
- Schmitt, E., M. Panvert, S. Blanquet, et Y. Mechulam. 1998. « Crystal structure of methionyl-tRNA<sup>fMet</sup> transformylase complexed with the initiator formyl-methionyl-tRNA<sup>fMet</sup> ». *EMBO J.* 17 (23): 6819- 26.
- Schmitt, E., M. Panvert, C. Lazennec-Schurdevin, P. D. Coueux, J. Perez, A. Thompson, et Y. Mechulam. 2012. « Structure of the Ternary Initiation Complex aIF2-GDPNP-Methionylated Initiator tRNA ». *Nat. Struct. Mol. Biol.* 19: 450- 54. doi:10.1038/nsmb.2259.
- Schreier, M. H., B. Erni, et T. Staehelin. 1977. « Initiation of Mammalian Protein Synthesis. I. Purification and Characterization of Seven Initiation Factors ». *Journal of Molecular Biology* 116 (4): 727- 53.
- Schreier, M. H., et T. Staehelin. 1973. « Initiation of Eukaryotic Protein Synthesis: (Met-tRNA<sup>fMet</sup>-40S Ribosome) Initiation Complex Catalysed by Purified Initiation Factors in the Absence of mRNA ». *Nat New Biol* 242 (115): 35- 38.
- Schrodinger, LLC. 2010. « The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.0 », mai.
- Schuwirth, Barbara S., Maria A. Borovinskaya, Cathy W. Hau, Wen Zhang, Antón Vila-Sanjurjo, James M. Holton, et Jamie H. Doudna. 2005. « Structures of the Bacterial Ribosome at 3.5 Å Resolution ». *Science (New York, N.Y.)* 310 (5749): 827- 34. doi:10.1126/science.1117230.
- Selmer, M., C. M. Dunham, F. V. th Murphy, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, et V. Ramakrishnan. 2006. « Structure of the 70S Ribosome Complexed with mRNA and tRNA ». *Science* 313 (5795): 1935- 42. doi:10.1126/science.1131127.
- Seong, B.L., et U.L. RajBhandary. 1987. « Escherichia coli formylmethionine tRNA: Mutations in GGG/CCC sequence conserved in anticodon stem of initiator tRNAs affect initiation of protein synthesis and conformation of anticodon loop. » *Proceedings of the National Academy of Sciences of the USA* 84: 334- 38.
- Sette, M., P. van Tilborg, R. Spurio, R. Kaptein, M. Paci, C. O. Gualerzi, et R. Boelens. 1997. « The structure of the translational initiation factor IF1 from E.coli contains an oligomer-binding motif ». *EMBO J.* 16 (6): 1436- 43.
- Shi, H., et P. B. Moore. 2000. « The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited ». *RNA* 6 (8): 1091- 1105.
- Shine, J., et L. Dalgarno. 1974. « The 3'-terminal Sequence of Escherichia Coli 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites ». *Proceedings of the National Academy of Sciences of the United States of America* 71 (4): 1342- 46.
- Simonetti, A., S. Marzi, A. G. Myasnikov, A. Fabbretti, M. Yusupov, C. O. Gualerzi, et B. P. Klaholz. 2008. « Structure of the 30S Translation Initiation Complex ». *Nature* 455 (7211): 416- 20. doi:10.1038/nature07192.
- Stringer, E. A., P. Sarkar, et U. Maitra. 1977. « Function of Initiation Factor 1 in the Binding and Release of Initiation Factor 2 from Ribosomal Initiation Complexes in Escherichia Coli ». *The Journal of Biological Chemistry* 252 (5): 1739- 44.

- Suddath, F.L., G.J. Quigley, A. Mc Pherson, D. Sneden, J.J. Kim, S.H. Kim, et A. Rich. 1974. « Three-dimensional structure of Yeast phenylalanine transfer RNA at 3.0 Å resolution. » *Nature* 248: 20- 24.
- Sundari, R., L. Stringer, L. Schulman, et U. Maitra. 1976. « Interaction of initiation factor 2 with initiator tRNA ». *J. Biol. Chem.* 251: 3338- 45.
- Thomas, A., H. Goumans, H. O. Voorma, et R. Benne. 1980. « The Mechanism of Action of Eukaryotic Initiation Factor 4C in Protein Synthesis ». *European Journal of Biochemistry / FEBS* 107 (1): 39- 45.
- Tomsic, J., L. A. Vitali, T. Daviter, A. Savelsbergh, R. Spurio, P. Striebeck, W. Wintermeyer, M. V. Rodnina, et C. O. Gualerzi. 2000. « Late Events of Translation Initiation in Bacteria: A Kinetic Analysis ». *EMBO J* 19 (9): 2127- 36. doi:10.1093/emboj/19.9.2127.
- Trachsel, H., B. Erni, M. H. Schreier, et T. Staehelin. 1977. « Initiation of Mammalian Protein Synthesis. II. The Assembly of the Initiation Complex with Purified Initiation Factors ». *Journal of Molecular Biology* 116 (4): 755- 67.
- Tsai, A., A. Petrov, R. A. Marshall, J. Korfach, S. Uemura, et J. D. Puglisi. 2012. « Heterogeneous pathways and timing of factor departure during translation initiation ». *Nature* 487 (7407): 390- 93. doi:10.1038/nature11172.
- Valásek, Leos, Klaus H. Nielsen, Fan Zhang, Christie A. Fekete, et Alan G. Hinnebusch. 2004. « Interactions of Eukaryotic Translation Initiation Factor 3 (eIF3) Subunit NIP1/c with eIF1 and eIF5 Promote Preinitiation Complex Assembly and Regulate Start Codon Selection ». *Molecular and Cellular Biology* 24 (21): 9437- 55. doi:10.1128/MCB.24.21.9437-9455.2004.
- Voigts-Hoffmann, F., S. Klinge, et N. Ban. 2012. « Structural Insights into Eukaryotic Ribosomes and the Initiation of Translation ». *Curr. Opin. Struct. Biol.* 22 (6): 768- 77. doi:10.1016/j.sbi.2012.07.010.
- Wei, C. L., S. E. MacMillan, et J. W. Hershey. 1995. « Protein Synthesis Initiation Factor eIF-1A Is a Moderately Abundant RNA-Binding Protein ». *The Journal of Biological Chemistry* 270 (11): 5764- 71.
- Weiel, J., et J. W. Hershey. 1982. « The Binding of Fluorescein-Labeled Protein Synthesis Initiation Factor 2 to Escherichia Coli 30 S Ribosomal Subunits Determined by Fluorescence Polarization ». *The Journal of Biological Chemistry* 257 (3): 1215- 20.
- Westhof, E., P. Dumas, et D. Moras. 1985. « Crystallographic Refinement of Yeast Aspartic Acid Transfer RNA ». *Journal of Molecular Biology* 184 (1): 119- 45.
- Williams, N. P., A. G. Hinnebusch, et T. F. Donahue. 1989. « Mutations in the Structural Genes for Eukaryotic Initiation Factors 2 Alpha and 2 Beta of Saccharomyces Cerevisiae Disrupt Translational Control of GCN4 mRNA ». *Proceedings of the National Academy of Sciences of the United States of America* 86 (19): 7515- 19.
- Wimberly, B. T., D. E. Brodersen, W. M. Clemons Jr., R. J. Morgan-Warren, A. P. Carter, C. Vornheim, T. Hartsch, et V. Ramakrishnan. 2000. « Structure of the 30S ribosomal subunit ». *Nature* 407 (6802): 327- 39.
- Wintermeyer, W., et C. Gualerzi. 1983. « Effect of Escherichia Coli Initiation Factors on the Kinetics of N-Acph-e-tRNAPhe Binding to 30S Ribosomal Subunits. A Fluorescence Stopped-Flow Study ». *Biochemistry* 22 (3): 690- 94.
- Woese, C. R., et G. E. Fox. 1977. « Phylogenetic Structure of the Prokaryotic Domain: The Primary Kingdoms ». *Proceedings of the National Academy of Sciences of the United States of America* 74 (11): 5088- 90.
- Yoon, H. J., et T. F. Donahue. 1992. « The Suil Suppressor Locus in Saccharomyces Cerevisiae Encodes a Translation Factor That Functions during tRNA(iMet) Recognition of the Start Codon ». *Mol Cell Biol* 12 (1): 248- 60.
- Yusupov, M. M., G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. Cate, et H. F. Noller. 2001. « Crystal structure of the ribosome at 5.5 Å resolution ». *Science* 292 (5518): 883- 96.

## Etude du complexe de démarrage de la traduction chez les archées

**Mots clés :** démarrage de la traduction ; archées ; toeprint ; anisotropie de fluorescence

Chez les archées, le démarrage de la traduction met en jeu un complexe macromoléculaire formé de la petite sous-unité du ribosome, d'un ARNm, d'un ARN de transfert initiateur méthionylé (Met-ARNt<sub>i</sub><sup>Met</sup>) et de trois facteurs de démarrage de la traduction (aIF1, aIF1A et aIF2). De manière intéressante, ces trois facteurs de démarrage ont chacun un orthologue eucaryote.

Les ARNt<sub>i</sub> archées et eucaryotes possèdent une paire de bases très conservée A<sub>1</sub>-U<sub>72</sub>, au sommet de la tige acceptrice. Cette paire de base a été montrée importante pour la discrimination des ARNt initiateurs et élongateurs. De plus, des travaux suggèrent l'importance de la géométrie de la paire A<sub>1</sub>-U<sub>72</sub> pour l'identité initiatrice de ces ARNts. Mon travail de thèse a consisté en la construction d'une souche d'*E. coli* utilisant comme seule source d'ARNt<sub>i</sub> un variant d'ARNt initiateur bactérien (ARNt<sub>f</sub><sup>Met</sup>) possédant une paire de base A<sub>1</sub>-U<sub>72</sub> (ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>). L'utilisation de cette souche nous a permis d'obtenir de grandes quantités d'ARNt<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> purifié. La structure cristallographique de cet ARNtA<sub>1</sub>-U<sub>72</sub> a pu être déterminée à 2.8 Å de résolution. Un arrangement inhabituel des bases A<sub>1</sub> et U<sub>72</sub> a été observé.

Tous les acteurs du démarrage de la traduction de l'archée *P. abyssi* étant disponibles au laboratoire, une étude du complexe de démarrage de la traduction archée par cryo-microscopie électronique a pu être réalisée. L'étude a permis d'identifier deux conformations de l'ARNt<sub>i</sub> dans le complexe de démarrage, ICO-P<sub>remote</sub> (5.3 Å de résolution) et IC1-P<sub>in</sub> (7.5 Å de résolution). Ces deux conformations permettent de proposer un modèle pour l'accommodation de l'ARNt initiateur lors de l'appariement au codon de démarrage.

Finalement, je me suis également intéressée au rôle du facteur aIF1. La disponibilité de structures 3D, de modèles d'assemblage, et d'alignements de séquences aIF1 d'archées ont permis de proposer des régions pouvant être impliquées dans la liaison au ribosome et/ou dans la formation du complexe de démarrage. Afin de pouvoir étudier l'implication de ces régions, j'ai mis au point une méthode d'étude de la liaison d'aIF1 à la petite sous-unité du ribosome par anisotropie de fluorescence. Cette étude met en évidence deux résidus basiques d'aIF1 impliqués dans la liaison au ribosome. D'autre part, les rôles d'aIF1 dans la sélection du codon de démarrage et dans la stabilisation de ce complexe sont étudiés par la méthode d'empreinte du ribosome ou toeprint.

## Study of the archaeal translation initiation complex

**Keywords :** translation initiation ; archaeal ; toeprint ; anisotropy of fluorescence

In archaea, initiation of translation involves a macromolecular complex containing the small subunit of the ribosome, mRNA, an initiator methionyl tRNA (Met-tRNA<sub>i</sub><sup>Met</sup>) and three initiation factors (aIF1, aIF1A and aIF2). Interestingly, each of three initiation factors has an ortholog in eukaryotes.

Archaeal and eukaryotic tRNA<sub>i</sub> have highly conserved bases A<sub>1</sub>-U<sub>72</sub>, at the extremity of the acceptor stem. This base pair was shown to be important to discriminate initiator tRNAs from elongator tRNAs. In addition, other studies suggest the importance of the geometry of the A<sub>1</sub>-U<sub>72</sub> pair for the initiator identity of those tRNAs. At the beginning of my thesis, no structural information was available to explain how the characteristics of the A<sub>1</sub>-U<sub>72</sub> pair were involved in the selection of the initiator tRNA. First, my thesis work involved the construction of an *E. coli* strain using as the sole source of tRNA<sub>i</sub>, a bacterial variant of tRNA initiator (tRNA<sub>f</sub><sup>Met</sup>) having a base pair A<sub>1</sub>-U<sub>72</sub> (tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>). The use of this strain allowed us to obtain large quantities of purified tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub>. In addition, the crystal structure of this tRNA<sub>f</sub><sup>Met</sup>A<sub>1</sub>-U<sub>72</sub> has been determined at 2.8 Å of resolution. An unusual arrangement of bases A<sub>1</sub> and U<sub>72</sub> was observed.

All archaeal translation initiation actors being available in the laboratory, a study of the archaeal translation initiation complex by cryo-electron microscopy was achieved. The study identified two conformations of the tRNA<sub>i</sub>. Analysis of both conformations, ICO-P<sub>REMOTE</sub> (5.3 Å resolution) and IC1-P<sub>IN</sub> (7.5 Å resolution), allowed us to propose a model for the accommodation of the initiator tRNA during start codon recognition.

Finally, I was also interested in the role of the aIF1 factor. Availability of 3D structures, assembly models and alignments of the archaeal aIF1 sequences allowed us to identify amino acids or regions that could be involved in ribosome binding and/or in the selection of initiator tRNA. In order to study the involvement of these regions, I have developed a method to study the binding of aIF1 to the small ribosomal subunit using fluorescence anisotropy. This study highlights two basic residues of aIF1 involved in binding to the ribosome. On the other hand, the roles of aIF1 in the selection of the start codon and in the stabilization of the initiation complex on the mRNA were studied by the ribosome footprint method also called « toeprint ».