UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE des Sciences de la Vie et de la Santé

Centre de Biologie Intégrative, IGBMC, UMR7104, Illkirch

THÈSE présentée par :

Heena KHATTER

Soutenue le : **18 Septembre 2014**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Biophysique et biologie structurale

(Biophysics and Structural biology)

Structure-function studies of human ribosome complexes

Etudes structure-fonction de complexes du ribosome humain

THÈSE dirigée par :
  Dr. KLAHOLZ Bruno P.   DR, CNRS, IGBMC, Illkirch.

RAPPORTEURS :
  Dr. JENNER Lasse B.   Associate professor, Aarhus University, Denmark.
  Dr. MECHULAM Yves   DR, École Polytechnique, Paris.

AUTRES MEMBRES DU JURY :
  Dr. WESTHOF Eric   Professor, CNRS, IBMC, Strasbourg.

Dedicated to my parents, Chander and Renu, for their unconditional love and support.

We are kept from our goal, not by obstacles, but by a clear path to a lesser goal.

-Quote 9.24

Bhagavad-gītā

Acknowledgements

My PhD thesis has been possible only because of constant support and participation of several people. Some of the people who helped me in accomplishing this task deserve the credit equally.

First and foremost, I would like to thank my thesis advisor Dr. Bruno Klaholz, for giving me this opportunity to work on the fascinating subject of human ribosomes. With his constant support and guidance, the past four years have been a great learning experience. Sincere thanks for being so patient with my amateur and overtly enthusiastic attitude. I am glad that you never underestimated my potential and helped me to become a better scientist.

I would take this opportunity to express my gratitude to all my Jury members, Dr. Lasse Jenner, Dr. Yves Mechulam, Dr. Eric Westhof and Dr. Albert Weixlbaumer who enthusiastically agreed to read my thesis and evaluate my work. I am grateful for your presence here.

I am thankful to the people at the cell culture service, especially Betty and Leslie, who were always cooperative for trying different HeLa culture conditions. Thanks for providing HeLa cells every time I needed them.

I am thankful to Stefano, who helped in setting-up the ribosome purification; he was indeed a great support in the first year. I am also thankful to Anne-Sophie, who worked with me for two years. It was a valuable experience for me, and it was her willingness to learn that made it such a pleasure to work together. Also, I would like to thank all the Klaholz team members, past and present, for their help and assistance during the past four years.

I am indebted to those who made working here at CEBGS and now CBI, a pure joy. Kareem, Sacha and Kundhavai for helping me learn crystallography and for all the useful advices in data processing, Judith and Marie, for being supportive all the time and for the much needed pep-talks, Alastair for all his nagging and bossing around, Pierre for his assistance at the crystallisation platform, Morgan for all the delicious cakes and help with translations, Anna, Karima, for being around in the lab and for the enthusiastic tea-time discussions. I am thankful to all of you for creating a fabulous work environment.

I am grateful to Alexey, Adam and Kundhavai for patiently correcting my thesis and bringing it to its presentable form.

I would also like to thank Adam for all the fruitful discussions and '5-min' chats. His contagious enthusiasm and high spirits made research a lot more motivating.

I would take this opportunity to thank my friends who have made this journey truly enjoyable and memorable. Kamal, Alexey, Dhiraj, Iskander, Thanuja, Anna, Karima, Grigory, Firas, Vladimir, Alastair, Irina. We met as strangers, but it's because of them that living in

Strasbourg has been so much fun. Also, Priyanka and Saumya, who have been just a phone call away to share my happiness and troubles alike.

Most importantly, I would like to thank my parents, my brother and my entire family for being my support and strength all the time, without which this work would never have been accomplished. Leaving the comfort of home and homeland was not easy for me and I understand it wasn't easy for my family as well. Sincere thanks for letting me follow my dreams. No matter in which part of the world I would be, I know that I can always rely on you all for encouragement.

Last, but not the least I would like to thank Sasha for teaching me the nuances of cryo-EM, for sharing creative ideas and his experience in this enthralling field of translation. He has been a pillar of support and motivation all throughout. I am indebted to him for being a constructive critic, for believing in me, for never letting me yield to the circumstances and keeping my spirits high in testing times.

It was a pleasure working here at IGBMC.

Summary

Protein synthesis in the cell is catalyzed by the ribosome and is regulated by protein factors that bind transiently to the ribosome during the different phases of translation-initiation, elongation, termination and recycling. My work focussed on structural and functional aspects of these huge (2-4 MDa) protein synthesising machines. When I started working on this project in October 2010, a lot of structural studies had been done on the prokaryotic ribosome, both by crystallography and by cryo electron microscopy (cryo-EM) of functional ribosome complexes with tRNA, mRNA, and protein factors (for example Chandramouli et al., 2008; Marzi et al., 2007; Spahn et al., 2004). However, human ribosomes were not studied to atomic resolution because of their particularly complex structure. Also, they are inherently difficult to prepare to high homogeneity, which is a key requisite for high-resolution structural work.

The aim of my project was two-fold:

1.   Human ribosome structure determination using X-ray crystallography and cryo-EM to provide insights into the specific mechanism of protein synthesis and regulation, like with respect to antibiotic side effects.

2.   Elucidate the molecular mechanism of translation termination by forming *in-vitro* termination complexes for cryo-EM analysis and co-crystallize the 2 eukaryotic release factors (eRF1 and eRF3).

I established the purification of homogenous 80S ribosomes from HeLa cells using sucrose density gradients and polyethylenglycol (PEG) precipitation. The ribosomes were characterized biophysically by Analytical Ultracentrifugation, Mass spectrometry, Multiangle Laser Light Scattering and cryo-EM imaging and 3D reconstruction.

Human 80S samples from different preparations were frozen in the hydrated state as thin vitreous ice films (Dubochet et al., 1988) and cryo-EM images were collected on the high-resolution in-house Tecnai Polara F-30 electron microscope. 3D density maps were calculated from the selected particles at a medium resolution using EMAN2 to address the ribosome conformation and the potential presence of endogenous factors or tRNAs.

Moreover, cryo-EM was used to screen samples with modified conditions to obtain samples with homogenous distribution on grid. With the aim of getting a high-resolution structure I worked on the crystallization of human ribosomes. I setup crystallization trials with well-characterized, homogenous 80S samples in drops as well as capillaries, using various screens like PEG/Ion, Index from Hampton Research; MPD suite, PEGs suite from Qiagen. Initial hits were obtained in capillaries with a few conditions; however, they did not diffract X-rays. These conditions were reproduced and optimized in sitting drops after a series of trials; since the principle of counter diffusion in capillaries is different from vapour diffusion used in sitting drops. Plate-like crystals were obtained which diffracted at SLS (Swiss Light Source) synchrotron up to 26 Å resolution. This part of the work is described in my article (Khatter et al., 2014) in Nucleic Acids Research, 2014, providing a very promising basis for future high-resolution work on the human ribosome.

In the recent past, there have been interesting advances in eukaryotic ribosome studies using crystallography (Ben-Shem et al., 2010, 2011; Rabl et al., 2010; Klinge et al., 2011) and cryo–EM of functional ribosome complexes with tRNA, mRNA, and protein factors (Taylor et al., 2012; Anger et al., 2013; Pallesen et al., 2013). Previously, our lab has worked on prokaryotic initiation and termination of translation. Now we wanted to focus on eukaryotic termination to reveal the interactions between the 2 eukaryotic release factors (eRF1 and eRF3). eRF1 (Class I release factor) indentifies the stop codon and binds to the translating ribosome, followed by eRF3 (Class II release factor) association and release of the nascent peptide chain.

The structures of the isolated eRF1 and eRF3, and of the eRF1/3 complex (missing the catalytic GTPase domain) has been determined already, but N terminus and G domain of human eRF3 have not been well studied. I cloned the eukaryotic release factors (eRF1 and eRF3) in suitable vectors for expression in E. coli. Both the proteins were expressed, purified to biochemical homogeneity and concentrated to 7-15 mg/ml. The interaction of the two proteins was confirmed using Microscale thermophoresis and dissociation constant (kd) was determined to be 150 nM consistent with values found in the literature. (yes, values were determined to be 200nM for full erf1/3 complex, in 2010) These factors were co-crystallized, and a few hits were obtained with protein crystals. During screening at SLS beamline (Villigen, Switzerland), it was found that out of all these crystals obtained with different

precipitants, only one condition with Lithium sulphate as precipitant gave diffraction up to 4 Å resolution. Data sets were collected for this crystal and a few others which diffracted up to 8 Å. These crystals, for the condition with Li2SO4, belong to space group P412121 with 2 molecules of eRF1 and 1 copy of eRF3 in the asymmetric unit, which could be because of eRF1 existing as a dimer in solution. Higher resolution data sets need to be obtained, but a first molecular replacement solution has been found.

The purified ribosomes from HeLa cells (Khatter et al., 2014) fulfil an essential requirement for forming functional complexes to investigate the missing links in translation like termination. The ribosomal termination complex was assembled using purified 80S, eRF1, eRF3, mRNA, tRNA (uncharged) and characterized using cryo-EM. For analysis, 80 000 particles were selected and structure refined by image processing, revealing that it contains the tRNA in the E (exit) site rather than in the P (peptidyl transferase) site. This could be due to misplacement of mRNA or complete absence of mRNA in this structure. Also, ribosomes exist in 2 major states, ratcheted and non-ratcheted, which is due to movement of the small subunit, leading to heterogeneity in the sample set. The data were analysed using the image processing softwares, EMAN2 and Relion with splitting the particle sets, to look for even a small population which might have tRNA in P site. This is being evaluated. This structure obtained will be analysed by fitting the structures determined by crystallography. It should provide accurate information on the functional specificity of eukaryotic ribosomes, with a prospect of developing specific antibiotics preferentially targeting the function of the prokaryotic ribosome.

Alongside, in collaboration with the group of Yves Mély at the faculty of pharmacy, Illkirch, we have been addressing the question whether interactions exist between the ribosome and the viral proteins GAG (viral polyprotein) and NCP7 (a constituent of GAG). HeLa cells transfected with GAG or NCP7, were lysed, spun briefly to remove cell debris and enveloped organelles (mitochondria and nuclei) and the supernatant was used for analysis. I performed 20-50% sucrose density gradients for polysome analysis and checked co-sedimentation of ribosomes with GAG. Western blot analysis revealed co-localization of GAG with polyribosome fractions. Even purified 80S ribosomes, when incubated with NCP7 peptide, co-precipitated, showing that NCP7 binds to the ribosome and could be involved in its own translational regulation.

Future Perspectives

My work on human ribosomes has established a new hybrid way of obtaining crystals for inherently challenging molecules, based on interplay between biochemistry, cryo-EM and crystallization assays. Sample visualisation for obtaining feedback on homogeneity and stability of sample is a novel aspect which is important to consider for the study of complex macromolecular assemblies in integrated structural biology. Here, it clearly highlights the advantage of teaming up two major structural techniques, which can be used in future for other biomolecules.

With the new in-house Titan Krios electron microscope installed recently better data sets can be collected like in the movie mode for achieving improved structures with less number of particles. High-resolution structure for termination complex will shed light on the exact mode of interaction of eRF1 and eRF3 with the peptidyl transferase center (PTC) and the decoding center, with the ultimate aim of providing insights into the molecular basis of stop codon recognition by the class-I release factors.

Résumé de thèse

Dans la cellule, la synthèse protéique est catalysée par le ribosome, et régulée par des facteurs protéiques qui interagissent de manière transitoire, pendant les différentes phases d'initiation, élongation, terminaison et recyclage. Mon travail porte sur l'étude des aspects fonctionnels et structuraux de cette énorme (2-4 MDa) machinerie de synthèse. Lorsque j'ai commencé à travailler sur ce projet en octobre 2010, plusieurs études structurales avaient été publiées, à la fois en cristallographie aux rayons-X et en cryo-microscopie électronique, du ribosome procaryote fonctionnel, seul ou en complexe avec des ARN de transfert (ARNt), ARN messagers (ARNm) ou facteurs protéiques (Chandramouli et al., 2008; Marzi et al., 2007; Spahn et al., 2004). Cependant, et en raison de sa structure particulièrement complexe, le ribosome humain n'avait alors pas été étudié à résolution atomique. De plus, sa purification homogène –un prérequis majeur pour l'étude à haute résolution– est tout particulièrement difficile à accomplir.

Le but de mon projet était double :

1. Résoudre la structure du ribosome humain par cristallographie aux rayons-X et cryo-microscopie électronique, afin de poser les bases structurales de la synthèse protéique et de sa régulation, par l'action d'antibiotiques par exemple.
2. Elucider le mécanisme moléculaire de la terminaison, en formant *in vitro* des complexes de terminaison pour l'étude par cryo-microscopie électronique, et en co-cristallisant les deux facteurs de terminaison eucaryotes (eRF1 et eRF3).

J'ai mis au point la purification de ribosomes 80S homogènes, à partir de cellules HeLa, en utilisant des gradients de densité de sucrose ainsi qu'une précipitation au polyéthylène glycol (PEG). Les ribosomes ont ensuite pu être caractérisé par plusieurs méthodes biophysiques, telles que l'ultracentrifugation analytique, la spectrométrie de masse, la diffusion de lumière-laser dynamique à plusieurs angles (MALLS) et la cryo-microscopie électronique avec reconstitution tridimensionnelle. Des échantillons de ribosome 80S de préparations différentes ont été congelés à l'état hydraté en fine couche de glace vitrifiée (Dubochet et al., 1988) et des images par cryo-microscopie électronique ont été collectées à haute résolution, sur le microscope Tecnai Polara F-30 installé dans

notre laboratoire. Des cartes de densité tridimensionnelles, à résolution moyenne, ont pu être calculées à partir des particules sélectionnées à l'aide du logiciel EMAN2, pour déterminer la conformation du ribosome et la présence potentielle de facteurs endogènes ou d'ARNt (Figure I).
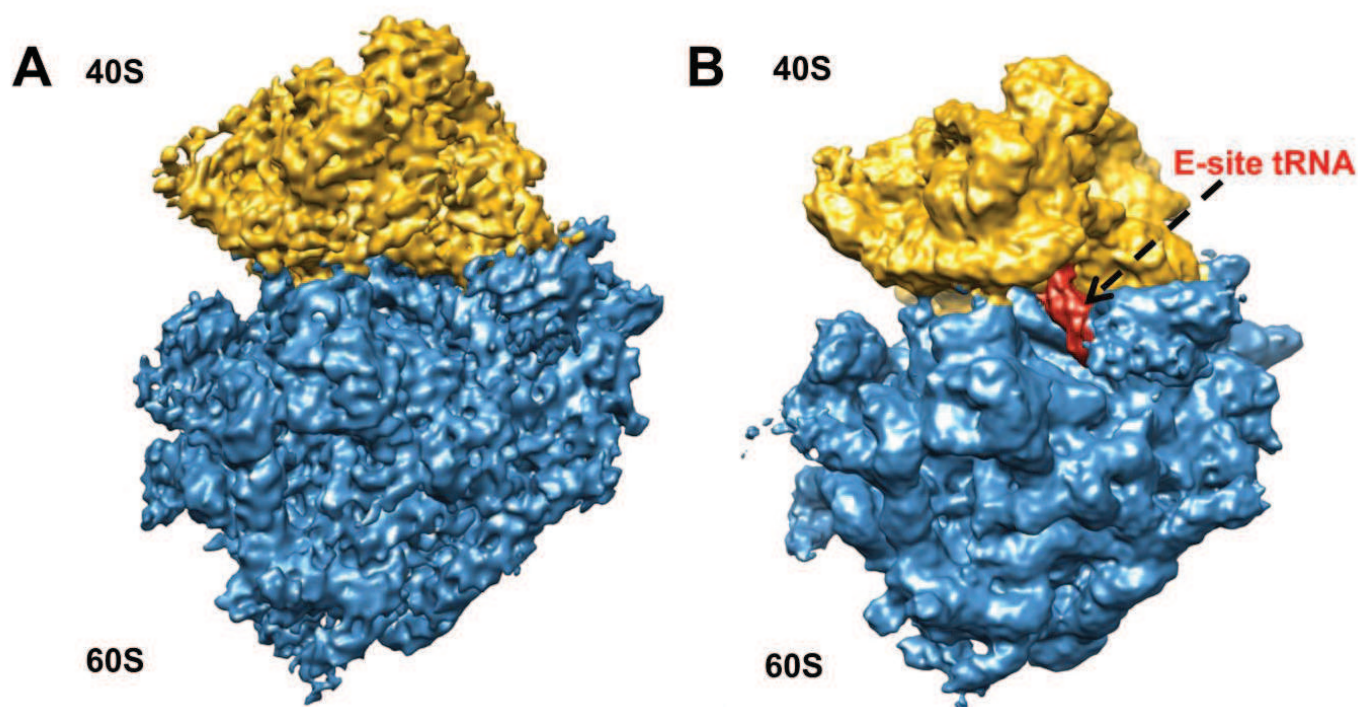

Figure I: (A) La structure en cryo-microscopie électronique du ribosome humain 80S vide et (B) avec un ARNt dans le site E ; les deux ont été purifiés à partir de cellules HeLa, à haute et faible concentration en KCl pendant le traitement à la puromycine, respectivement ; code couleur : 40S en jaune, 60S en bleu, ARNt en rouge.

De plus, la cryo-microscopie électronique a été mise à profit pour cribler des échantillons dans différentes conditions, afin d'obtenir une distribution homogène sur les grilles. Dans le but d'obtenir une structure à résolution atomique, j'ai également travaillé à la cristallisation du ribosome humain. J'ai mis au point des essais de cristallisation à partir d'échantillons de ribosomes purs et préalablement caractérisés par les différents moyens énoncés précédemment, en gouttes assises, suspendues, mais également en capillaires, avec différents kits commerciaux de criblage (PEG/Ion et Index de Hampton Research, MPD suite et PEGs suite de Qiagen). Des premiers cristaux ont été obtenus en capillaires dans plusieurs conditions. Cependant, ils ne diffractaient pas les rayons-X, même sur les lignes intenses des synchrotrons. Le principe de contre-diffusion en capillaire étant différent de

celui de diffusion de vapeur en gouttes assises, ces conditions ont dû être re-optimisées pour la mise en place de gouttes assises. Des cristaux en forme de plaques ont été obtenus, et leur diffraction à la source synchrotron SLS (Swiss Light Source) a atteint 26 Å de résolution, illustrant la première obtention de cristaux du ribosome humain capables de diffracter les rayons-X (Figure II). Cette partie de mon travail a fait l'objet d'un article (Khatter et al., 2014) dans *Nucleic Acids Research, 2014*, fournissant une base très prometteuse pour les futurs travaux à haute résolution sur le ribosome humain.
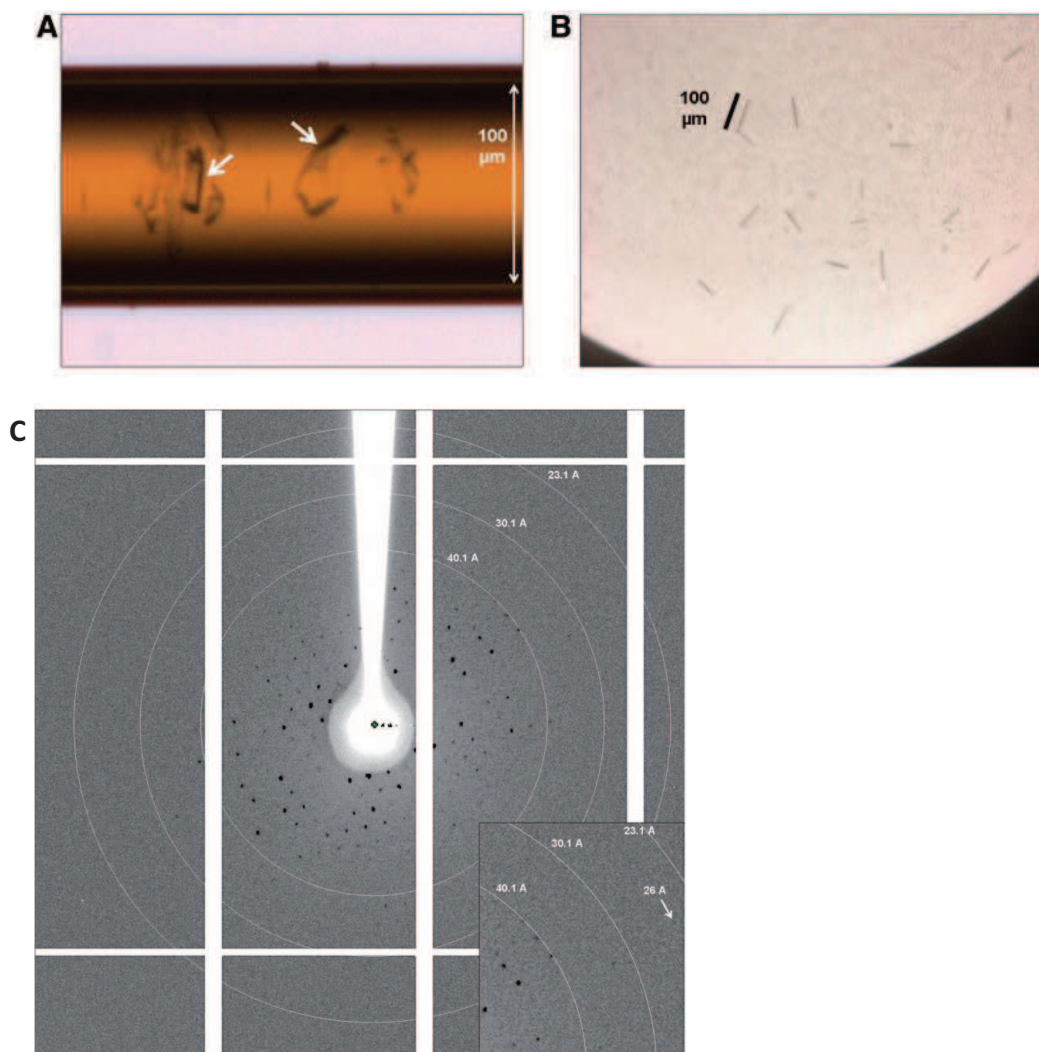

Figure II: (A) Premiers cristaux du ribosome humain 80S, obtenus dans les capillaires d'une plaque de cristallisation CrystalHarp™. (B) Cristaux en forme de plaque reproduits en goutte assise, diffractant à 26 Å. La plupart sont visualisés sur leur tranche et donnent ainsi l'impression de baguettes. (C) Le spectre de diffraction montre un réseau réciproque complet avec des paramètres de maille d'environ a = 406 Å, b = 785 Å et c = 977 Å. Les cercles de résolutions affichés sont à 23, 30 et 40 Å. Le cliché montre des tâches de diffraction jusqu'à 26 Å de résolution.

Récemment, des avancées intéressantes dans le domaine des ribosomes eucaryotes ont vu le jour, grâce à la cristallographie aux rayons-X (Ben-Shem et al., 2010, 2011; Rabl et al., 2010; Klinge et al., 2011) ou à la cryo-microscopie électronique (Taylor et al., 2012; Anger et al., 2013; Pallesen et al., 2013). Par le passé, notre laboratoire s'est concentré principalement sur l'initiation et la terminaison de la traduction procaryote. Nous avons donc souhaité faire un pas en avant et étudier la terminaison de la traduction chez les eucaryotes, et notamment l'interaction observée entre les deux facteurs de terminaison eRF1 et eRF3. eRF1 (facteur de terminaison de classe 1) est capable d'identifier les codons stop et de se lier au ribosome en phase d'élongation. eRF3 (facteur de terminaison de classe 2) est alors recruté, et permet la libération de la chaîne peptidique néosynthétisée.

Les structures d'eRF1 et eRF3 isolés, ainsi que du complexe eRF1/3 (sans le domaine catalytique GTPase) ont déjà été publiées, mais nous manquons d'informations concernant l'extrémité amino-terminale ainsi que le domaine G de la protéine eRF3 humaine (Figure III). J'ai donc entrepris de cloner en vecteurs bactériens les deux gènes erf1 et erf3. Après expression en bactérie *E. coli*, j'ai pu isoler les deux protéines et obtenir des concentrations de protéine pure allant de 7 à 15 mg/mL. L'interaction entre les deux facteurs a été mesurée par thermophorèse à micro-échelle et la constante de dissociation (k_d) a été estimée à 150 nM. Ce résultat corrobore les valeurs décrites dans la littérature. Ces facteurs ont alors été co-cristallisés, et quelques cristaux ont pu être obtenus dans différentes conditions. L'ensemble des cristaux a alors été testé à la source synchrotron SLS avec une diffraction atteignant les 8 Å de résolution et des premiers jeux de données ont été enregistrés. Apres optimisation de tous les cristaux obtenus avec différents agents précipitants, seule une condition avec du sulfate de lithium a permis d'obtenir une diffraction à 4 Å de résolution. Un jeu de données a donc été collecté pour ce cristal qui appartient au groupe d'espace $P4_12_12_1$, avec deux molécules d'eRF1 et une molécule d'eRF3 par unité asymétrique. Ceci peut s'expliquer par l'observation faite de dimères d'eRF1 en solution. Des jeux de données à plus haute résolution sont nécessaires pour obtenir une structure plus détaillée, mais un premier remplacement moléculaire a pu être effectué sur ces jeux à moyenne résolution.
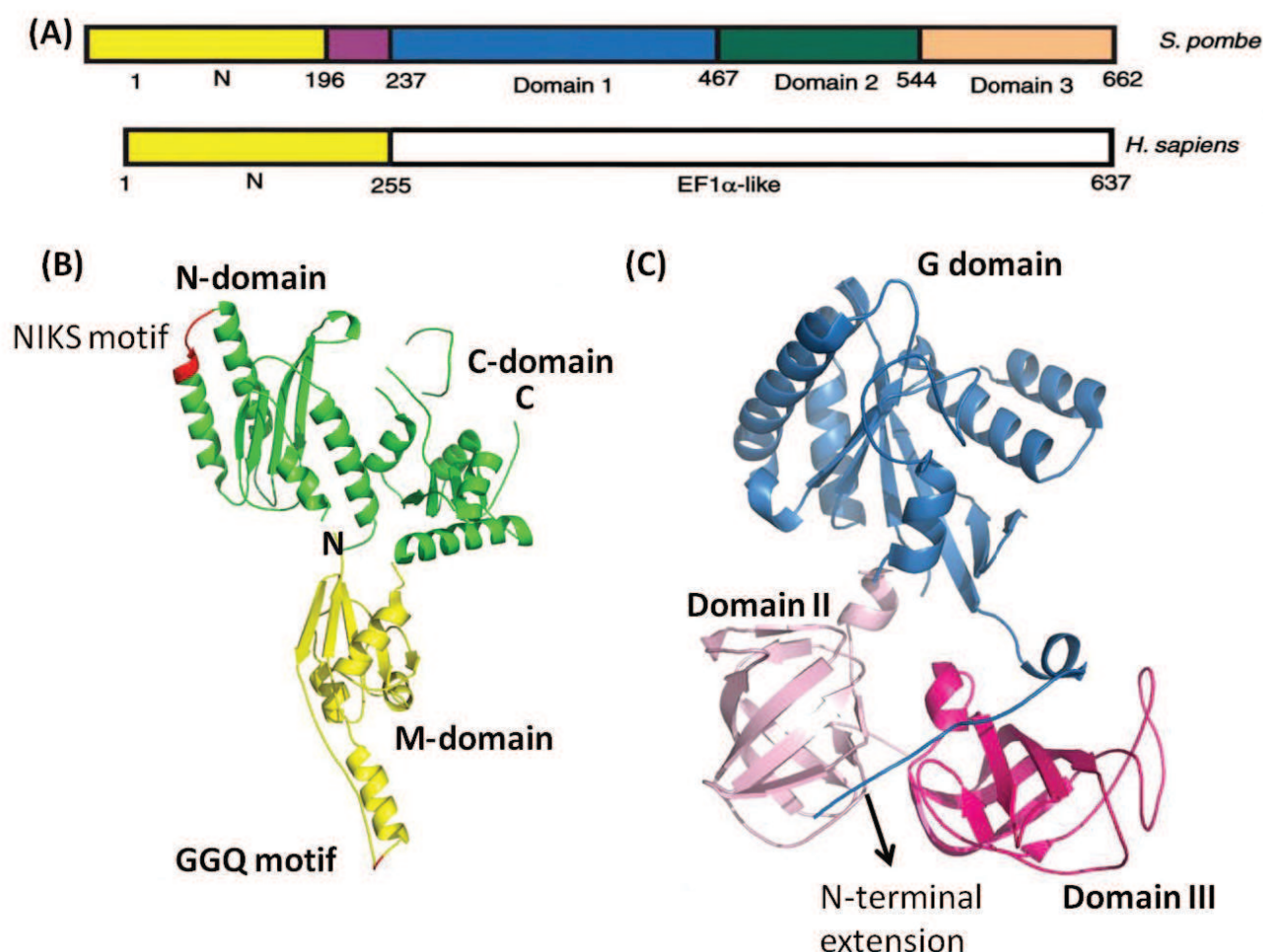
Figure III: (A) Représentation schématique des protéines eRF3 de *S. pombe* et humaine. (B) La structure cristallographique de eRF1 montre les motifs catalytiquement actifs NIKS et GGQ, marqués en rouge dans le domaine N (vert) et M (jaune), respectivement. Code d'accession PDB : 1DT9. (C) La structure cristallographique de eRF3 sans le domaine amino-terminal, montrant les domaines 1, 2 et 3 en bleu, rose et magenta, respectivement. Code d'accession PDB : 1R5B.

Les ribosomes purifiés à partir de cellules HeLa (Khatter et al., 2014) remplissent les conditions essentielles à la formation de complexes fonctionnels, afin d'étudier l'étape encore mal comprise de la terminaison. Le complexe de terminaison ribosomal a été assemblé *in vitro*, en mélangeant du ribosome humain 80S purifié, eRF1, eRF3, un ARNm et un ARNt non chargé, et ce complexe a pu être caractérisé par cryo-microscopie électronique. Pour son analyse, 80000 particules ont été sélectionnées, et la structure, affinée par traitement d'images, a montré que l'ARNt s'était positionné dans le site de sortie (E) plutôt que dans le site peptidyl-transférase (P) (Figure IV). Ceci peut être dû à un mauvais ancrage de l'ARNm, voire son absence du complexe. De plus, les ribosomes existent

majoritairement sous deux états, « *ratcheted* » et « *non-ratcheted* », dus à une légère rotation de la petite sous-unité, provoquant ainsi une hétérogénéité de l'échantillon. Les données ont été traitées à l'aide des logiciels EMAN2 et Relion, en subdivisant les jeux de données, afin de mettre en évidence une population, aussi minoritaire soit-elle, de ribosomes présentant l'ARNt dans le site P. Ceci est en cours de traitement et la structure obtenue le cas échéant sera analysée par intégration des structures cristallographiques disponibles. Ceci devrait fournir des informations précises quant à la spécificité fonctionnelle des ribosomes eucaryotes, avec la perspective de pouvoir développer des antibiotiques spécifiques, ciblant préférentiellement les ribosomes procaryotes.
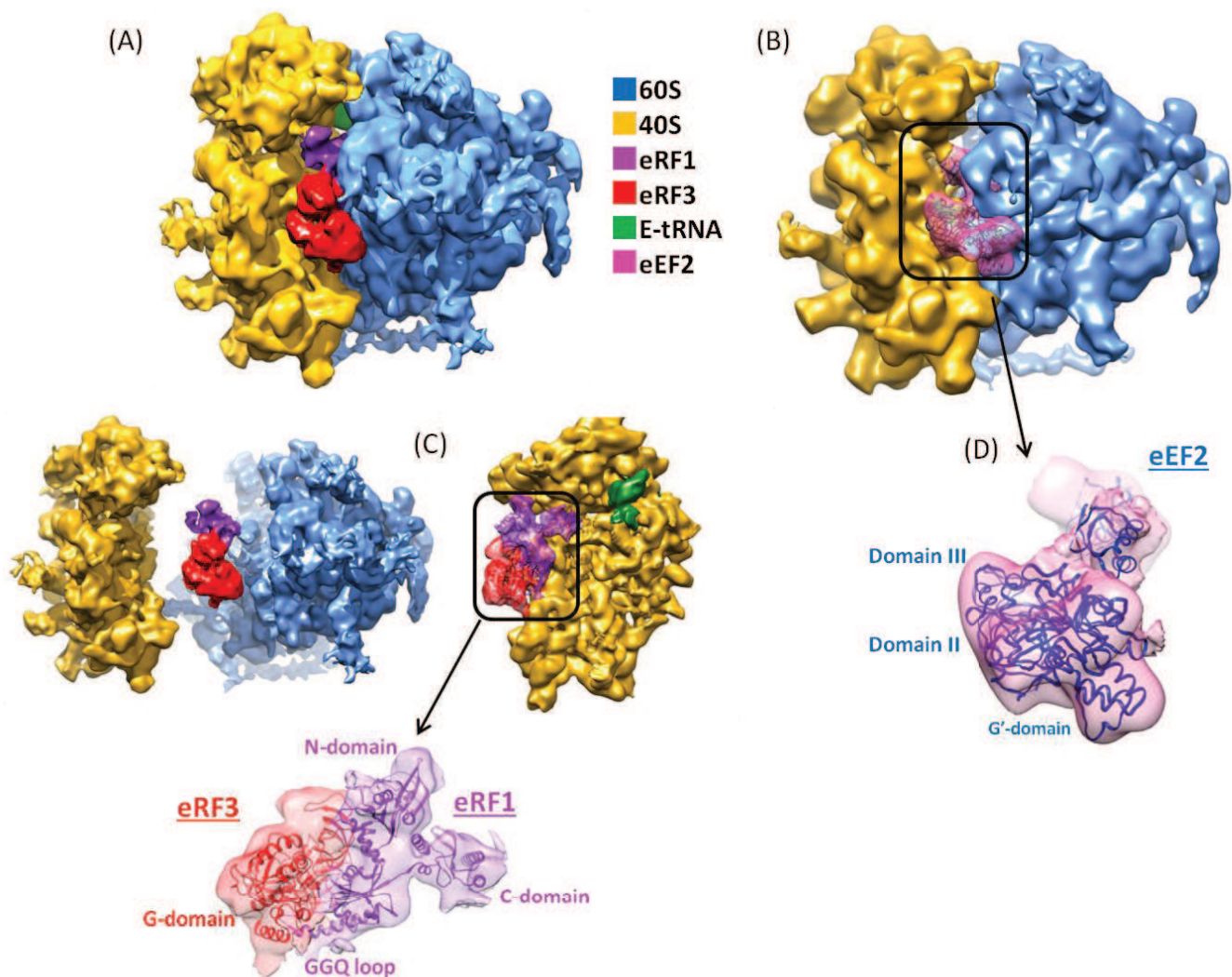

Figure IV: Le complexe de terminaison reconstitué *in vitro*, avec une courte séquence d'ARNm. (A) Vue de côté du complexe, montrant le facteur de terminaison dans son site de liaison, eRF1 en violet, eRF3 en rouge et l'ARNt dans le site E en vert. (B) Vue de côté du complexe montrant eEF2. (C) La séparation des densités électroniques montre les deux sous unités du ribosome ainsi que les deux facteurs de terminaison. La structure cristallographique du complexe eRF1-eRF3 superposée (code d'accession PDB : 3J5Y) montre que le motif GGQ s'oriente et pointe à l'opposé du PTC

Parallèlement, en collaboration avec le groupe d'Yves Mély à la faculté de pharmacie d'Illkirch, nous avons voulu répondre à l'hypothèse d'interactions entre les ribosomes et les protéines virales GAG (polyprotéine virale) et NCP7 (un constituant de GAG). Des cellules HeLa transfectées avec GAG ou NCP7, ont été lysées, brièvement centrifugées pour supprimer les débris cellulaires et les organites à enveloppes (mitochondries et noyaux) ainsi que le surnageant ont été utilisés pour l'étude. J'ai effectué des gradients de densité de sucrose de 20 à 50 % pour l'analyse de polysomes et vérifié la co-sédimentation des ribosomes avec GAG. Les analyses par western blot ont révélé la co-localisation de la protéine GAG avec les fractions de polyribosomes. De plus, les ribosomes 80S purifiés, et incubés avec un NCP7, ont pu être co-précipités, suggérant que NCP7 se lie au ribosome et pourrait être impliqué dans sa propre régulation traductionnelle.

Perspectives futures

Mon travail sur le ribosome humain a permis d'établir une nouvelle méthode synergique pour l'obtention de cristaux dans le cadre de molécules particulièrement difficiles, basée sur l'interface entre la biochimie, la cryo-microscopie électronique et les essais de cristallisation. La visualisation des échantillons et des particules des complexes macromoléculaire le constituant afin d'obtenir un aperçu de leur homogénéité et stabilité est un nouvel aspect qu'il est important de prendre en compte pour l'étude de complexes macromoléculaires en biologie structurale intégrative. Dans le cas du ribosome humain, cette méthode montre très clairement ses avantages, et pourra être utilisée à l'avenir dans le cadre d'autres biomolécules. Grâce au tout nouveau Titan Krios installé récemment au Centre de Biologie Intégrative (CBI) à l'IGBMC, de meilleurs jeux de données pourront être collectés, notamment en utilisant un mode d'enregistrement continu, pendant lequel le détecteur enregistre jusqu'à dix images par seconde sous le faisceau d'électron. Enfin, la structure à haute résolution du complexe de terminaison permettra de mettre en lumière le mécanisme précis d'interaction d'eRF1 et eRF3 avec le centre peptidyl-transférase et le centre de décodage du ribosome, avec le but ultime d'apporter la réponse à la reconnaissance moléculaire des codons stop par les facteurs de terminaison.

Publications

1. Jean-François Ménétret, <u>Heena Khatter</u>, Angelita Simonetti, Igor Orlov, Alexander G. Myasnikov, Vidhya KV, Sankar Manicka, Morgan Torchy, Kareem Mohideen, Anne-Sophie Humm, Isabelle Hazemann, Alexandre Urzhumtsev, Bruno P. Klaholz. **Integrative structure-function analysis of large nucleoprotein complexes.** RNA Structure and Folding, edited by D. Klostermeier & C. Hammann, 01/2013: chapter 10.2: pages 250-259; de Gruyter., ISBN: 1437-4315.

2. <u>Khatter, H.</u>, Myasnikov, A. G., Mastio, L., Billas, I. M. L., Birck, C., Stella, S., and Klaholz, B. P. (2014) **Purification, characterization and crystallization of the human 80S ribosome**, *Nucleic Acids Res., **2014**, 42(6), 1-11; doi:10.1093/nar/gkt1404.*

Oral and Poster Presentations

1. Integrated structural biology approach for human ribosome studies. IGBMC Internal seminar. 27th March 2014. **Heena Khatter.** <u>Oral presentation</u>

2. Integrated structural biology approach for ribosome studies. **Heena Khatter.** School of Life Sciences, Jawaharlal Nehru University, India. 7th November 2013. <u>Oral presentation</u> as an ambassador for IGBMC PhD program.

3. Characterization, cryo-EM analysis and crystallization of the human 80S ribosome. **Heena Khatter.** "27th Rhine-Knee Regional Meeting on Biocrystallography", 25-27 September 2013, Schluchsee, Germany. <u>Oral presentation</u>.

4. Purification, characterization and crystallization of the human 80S ribosome. **Heena Khatter,** Alexander Myasnikov, Stefano Stella and Bruno Klaholz. "RAMC 2013". 08-11 September 2013, Bischoffsheim, France. <u>Poster</u>.

5. Structure-function study of the human ribosome. **Heena Khatter**, Alexander Myasnikov, Stefano Stella, Catherine Birck, Isabelle Billas and Bruno Klaholz. "CNRS - Jacques Monod Conference The translating ribosome: towards mature proteins" 2-6 June, 2012 Roscoff, France. <u>Poster</u>.

Table of Contents

Table of Figures

Abbreviations

Amp Ampicillin

ATP Adenosine triphosphate

cryo-EM Cryo electron microscopy

CV Column Volume

DLS Dynamic light scattering

DTT Dithiothreitol

EDTA Ethylene diamine tetra acetate

eRF1 eukaryotic Release Factor 1, termination factor Class I

eRF3 eucaryotic Release Factor 3, termination factor Class II

ESI-TOF Electrospray ionization time-of-flight

FPLC Fast Pressure liquid Chromatography

g gram

GDP Guanosine-diphosphate

GMPPCP 5'-Guanosyl-methylene-triphosphate

GTP Guanosine- triphosphate

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IPTG Isopropyl β-D-1-thiogalactopyranoside

ITC Isothermal Titration Calorimetry

kDa Kilodaltons

LB Luria Bertani

mM Millimolar

MST Microscale thermophoresis

NMD Nonsense Mediated-mRNA Decay

Ni- NTA Nickel-Nitriloacetic acid

PABP PolyA binding protein

PI Protease inhibitor

SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoresis

pI Isoleectric point of protein

βME β-mercaptoethanol

1. Introduction

1.1 Overview of translation

The genetic information in a cell is stored in the DNA and it must be "read" to synthesise proteins. The double stranded DNA acts as a template, during transcription, to generate RNA. The RNA, synthesised by RNA polymerase (RNA pol), is required for the next step of translation to manufacture proteins. These two processes of transcription and translation comprise the central dogma in molecular biology. Three types of RNA are synthesised by their respective RNA polymerases (RNA pol). RNA pol II synthesises messenger RNA (mRNA). This mRNA dictates polypeptide sequence to be created by ribosome, as it is coded so that every three nucleotides of mRNA correspond to an amino acid. Transfer RNA (tRNA), produced by RNA pol III, carries specific amino acids to be incorporated in the growing peptide chain. And lastly, ribosomal RNA (rRNA), transcribed by RNA pol I, forms the catalytic component of the massive macromolecular machine called ribosomes.

Ribosomes are composed of two subunits, each possessing rRNA and protein components. Both subunits harbour specific centres to achieve this humongous and highly regulated task of polypeptide formation. The peptidyl transferase centre (PTC), where the peptide bond is formed; GTPase associated centre (GAC) where several GTP-binding factors hydrolyse GTP to provide a kinetic control mechanism for this process and the peptide exit tunnel which forms a conduit for nascent peptide chain passage; are present in the large subunit (60S/50S). The small (40S/30S) subunit possesses the mRNA channel for associating with mRNA and the decoding site where mRNA is recognised by tRNA.

The tRNA binding on the ribosome is compartmentalised into three (aminoacyl (A), peptidyl (P), and exit (E)) sites depending upon the state of the tRNA bound. The A-site binds the tRNA charged with an amino acid while the P-site tRNA carries the growing peptide chain and the E-site has the uncharged tRNA, ready to exit from the translation machinery. Each of these sites is partly present in both the subunits, which build upon assembly of the full ribosome. The presence of tRNA at each of these sites is annotated as A/A, P/P and E/E, with the first symbol denoting to the contact with the small subunit and second referring to the large subunit.
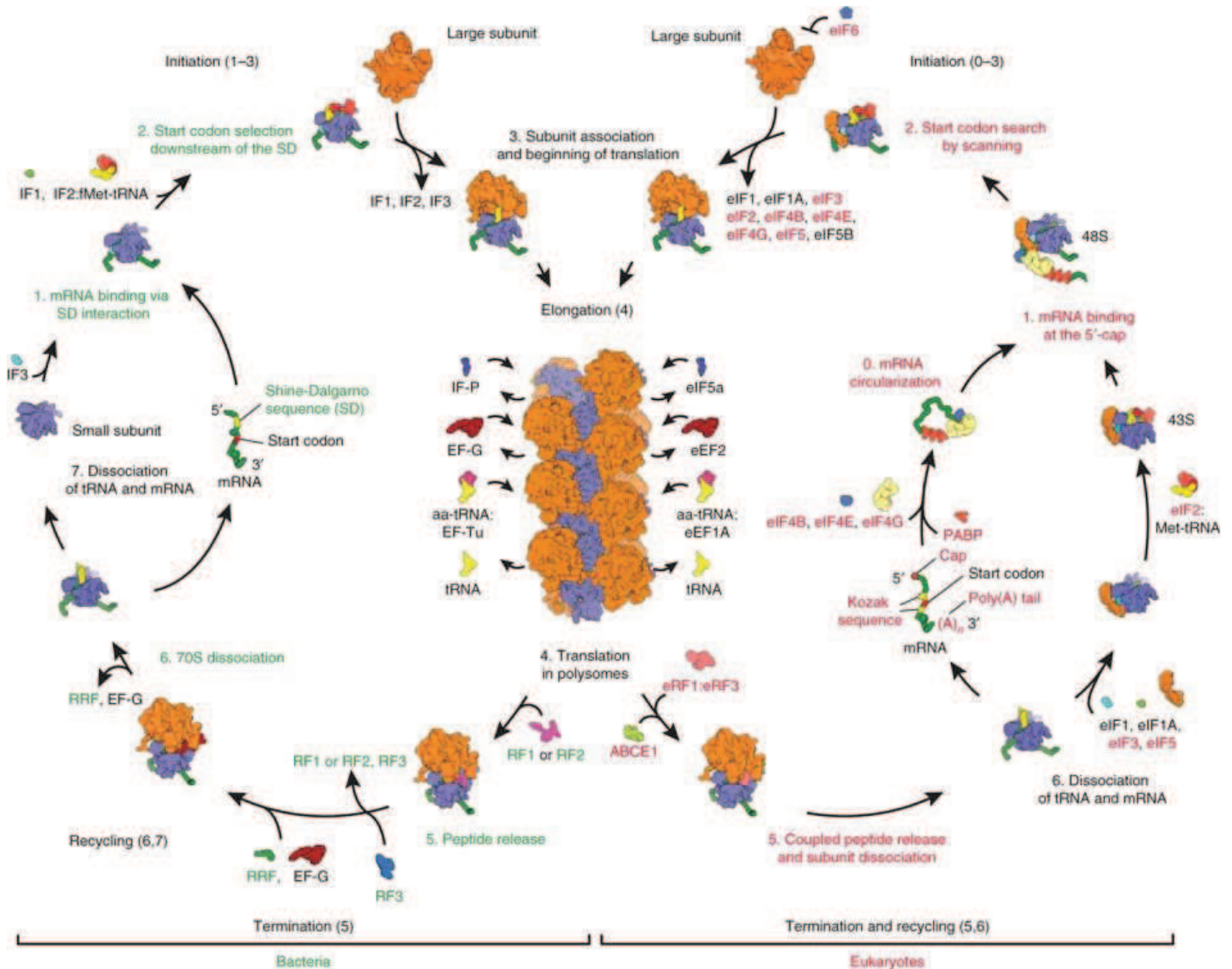
Figure 1: Scheme depicting translation cycle. In bacteria (on the left) and in eukaryotes (on the right). Adapted from Melnikov et. al., 2012.

Protein synthesis is an intricate process that combines high speed with high fidelity (Green & Noller, 1997). It is broadly categorised into four stages, initiation, elongation, termination and recycling. Initiation is the rate-limiting step as it requires start codon recognition by the small subunit (Fig. 1). Once the ribosome has assembled at the start site with the AUG start codon in the P-site, the charged aminoacylated tRNA is delivered by an elongation factor and there begins the next phase of translational elongation. It involves aminoacylated tRNA selection, peptide bond formation (Rodnina et al., 2006), tRNA-mRNA translocation (Frank et al., 2007; Spirin, 2009) in a repeated fashion until a stop codon is encountered in the A-site. The stop codon (UAG, UGA or UAA) recruits release factor instead

of tRNA which allows the release of the newly synthesised nascent peptide chain (Youngman et al., 2008). Finally, the 2 subunits are dissociated in the recycling step (Franckenberg et al., 2012) making them available for the next translation cycle. This basic mechanism of the protein synthesis is conserved across species, but there are stark differences in regulation and certain steps like initiation, associated with the higher complexity of life.

1.2 Key players in translation

The translation machinery requires RNA and protein components. RNA is synthesised as a single stranded molecule which can form secondary structures by folding over and forming hairpin loops. These structures are stabilised by intramolecular hydrogen bonds formed between the complementary bases (A:U, G:C), as observed for all three types of RNA.

mRNA is primarily composed of coding sequences that carry genetic information for sequence of protein to be synthesised. In addition, there are stretches of non-coding or untranslated regions at the 5'- and 3'- ends. Prokaryotic mRNA is ready for translation immediately after transcription. However, in eukaryotes its post-transcriptional processing differs from prokaryotes. The eukaryotic mRNA must undergo 5'-cap addition, splicing and 3'-polyadenylation (Fig. 2A). The 5' cap is a modified guanine (7-methylguanosine) residue, added co-transcriptionally to the first nucleotide with 5'-5'-triphosphate bond, which is essential for ribosome recruitment. Furthermore, as soon as mRNA is completely transcribed, about 250 adenosine residues are added to the 3' end to allow export of the mRNA from the nucleus. 5' cap addition and polyadenylation also ensure that mRNA is not degraded in the cell by nucleolytic enzymes. Finally, it might contain some non-coding stretches (introns) in the open reading frame which are cleaved off from the pre-mRNA, during splicing. The protein-coding sequences are then joined, completing the processing phase.

tRNA acts as an adaptor molecule linking the nucleotide sequence to the amino acid. The single strand of tRNA loops back on itself to form a "clover leaf" secondary structure and compacts further to form 3D L-shaped structure (Holley, 1965; Holley et al., 1965) (Fig. 2B). Each of the 3 loops of tRNA has a structure-function association. The anticodon loop

possesses the 3 nucleotides that correspond directly with its specific mRNA codon. The D-loop and the TψC loops are signified by modified uracil bases, dihydrouridine and pseudouridine, respectively (Dudock et al., 1969). The 3' end has an unpaired CCA sequence, which is added by CCA-adding enzymes (tRNA nucleotidyl transferases) during maturation (Deutscher and Ni, 1982). This CCA end is a prerequisite for aminoacylation and is recognised by the enzyme, aminoacyl tRNA synthetase which charges the tRNA with its amino acid (Xiong and Steitz, 2006). For each amino acid to be linked to the tRNA, a single aminoacyl tRNA synthetase exists in a cell.
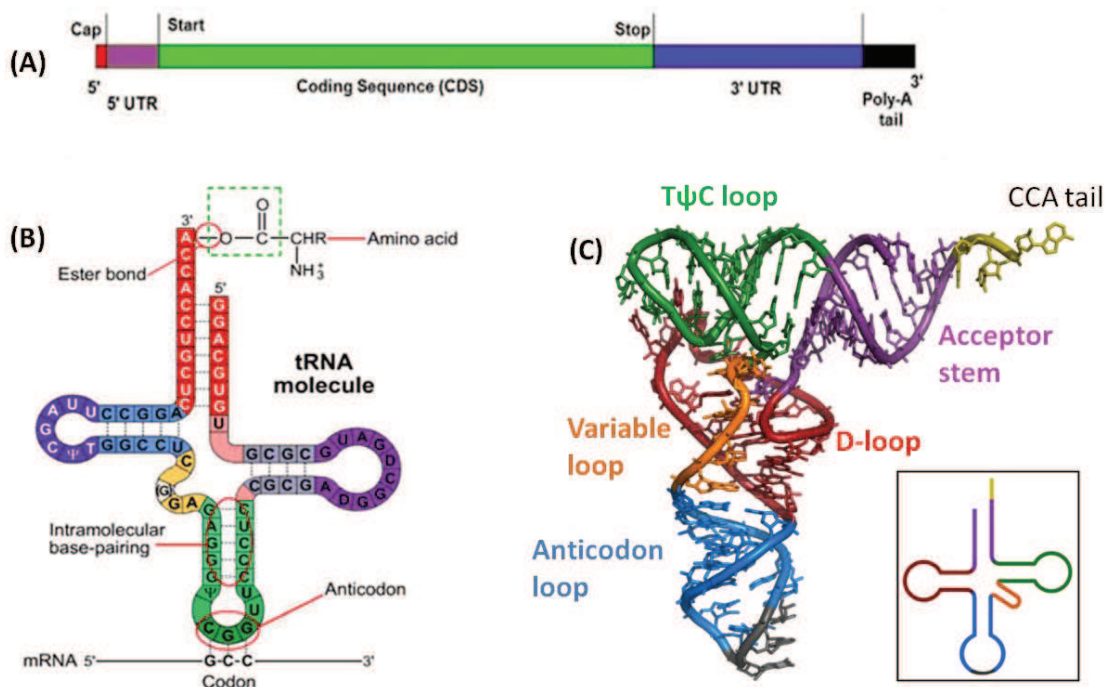
Figure 2: mRNa and tRNA. (A) Pictorial representation of eukaryotic mRNA. tRNA structure (B) 2D (C) 3D L –shaped, the inset shows the colour coding for different loops.

In prokaryotes, the fMet-tRNA charged with (N-formyl methionine) and in eukaryotes the Met-tRNA charged with methionine are the first residues to be delivered in synthesis of protein coded by AUG start codon on the mRNA (Marcus et al., 1970).

Factors: Translation is regulated by protein factors that bind transiently to the ribosome during the different phases (Table 1). Initiation is regulated by only three factors in prokaryotes (IF1, IF2 and IF3) (Jackson et al., 2010; Myasnikov et al., 2009) as compared to ten factors in eukaryotes. Each of these prokaryotic factors have eukaryotic equivalents, eIF1A, eIF5B and eIF1, performing similar functions (Eiler et al., 2013) (Table 1). eIF1 and

eIF1A are recruited during the recycling phase to the small ribosomal subunit. They allow the appropriate translation initiation site and the start codon selection during scanning on the mRNA (Martin-Marcos et al., 2011). eIF5B accelerates 60S-40S subunit joining. On the other hand, eIF6 is a unique anti association factor, which binds to the 60S and prevents its joining to the 40S. eIF2B promotes GDP-GTP exchange on eIF2. The rest of the eukaryotic initiation factors are majorly involved in either mRNA recruitment (eIF4, eIF3 and PABP) or assist in delivering Met-tRNAi to the 40S subunit (eIF2 and eIF5) apart from performing other important functions. eIF4A, a helicase, forms a component of eIF4F and allows unwinding of the 5' region of mRNA and attachment of the pre-initiation complex to mRNA; eIF4B like eIF4H is an RNA binding protein that enhances the helicase activity of eIF4A (Hinnebusch, 2011). These factors make initiation in eukaryotes a complicated and a highly regulated process.

Table 1: Canonical protein factors involved in translation. The factors performing similar functions are present in the same line.

Step in translation	Factors involved in prokaryotic translation	Factors involved in eukaryotic translation
Initiation	IF1 IF2 IF3	eIF1A eIF5B eIF1 eIF2 eIF2B eIF3 eIF4 eIF5 eIF6 PABP
Elongation	EF-Tu EF-G	eEF1A eEF2
Termination	RF1 and RF2 RF3	eRF1 eRF3
Recycling	RRF EF-G	ABCE1 Pelota eRF1 eIF3 eIF1A eIF1

During elongation eEF1A (EF-Tu in bacteria) delivers the aminoacyl tRNA to the A-site while eEF2 (EF-G in bacteria) provides directionality to the tRNA movement from A to P site

and prevents tRNA from back translocation (Voorhees and Ramakrishnan, 2013). Also, each of these steps has at least one GTPase-associated factor (eIF2, eEF1A, eRF3) which, most likely, fulfils the energy requirement for translation. The role of release and recycling factors is discussed in section III.

<u>Ribosome</u> forms the hub of translation. It is a highly versatile machine which allows recruitment of all these components in a stepwise manner and actively participates in all the steps. In its fully assembled 70S or 80S form it coordinates two major activities of mRNA decoding and peptide bond formation. Below, its structure is discussed in detail.

1.3 The prokaryotic ribosomal structure

Ribosomes are about 20nm in diameter and range from 2-4 MDa in molecular weight, with protein and rRNA components present in 2:1 ratio in prokaryotes as well as eukaryotes. The two subunits have disproportionate component distribution. The small ribosomal subunit possesses a single rRNA chain while the large subunit possesses two rRNA chains in prokaryotes and three rRNA chains in eukaryotes. The overall composition of ribosomes was determined by biochemical studies. However, providing insights into the structural aspects and how these components organise themselves to carry out translation, the two major structural techniques were used cryo-electron microscopy and crystallography.

Like biochemical studies, structural studies on ribosome complexes also involve several technical hurdles and challenges that needed to be overcome. In the context of crystallography, starting from obtaining a homogenous sample to solving the structure, every step had to be optimised to obtain structure of this huge MDa machine. Due to the presence of several flexible components and rRNA that is highly prone to degradation, obtaining crystals of this huge complex was and continues to be a major challenge. Moreover, the lifetime of crystals in an X-ray beam and collecting useful data sets are another issue. The advent of cryo-crystallography technique and synchrotron facilities for X-rays facilitated the acquisition of complete datasets. Interpreting this diffraction data represents another challenge in terms of model building.

For solving any structure by X-ray diffraction a "phase problem" is encountered. The detectors record the intensity of diffracting x-rays, but miss out on recording the phases. Adding heavy atoms to the crystal, allows determining the phases with respect to the relative position of the atoms in the crystal. But the diffraction power of single heavy atoms is too low compared to the size of the ribosome. Special heavy metal clusters like W18 containing about 2000 atoms came to the rescue as they could be detected in Patterson difference maps (Ban et al., 1998; Ban et al., 2000). In addition, the low resolution electron density maps from electron microscopy were then used to obtain phases at low resolution and further, to localise heavy atom sites and perform phase extensions towards high resolution (Cate et al., 1999; Yusupov et al., 2001).
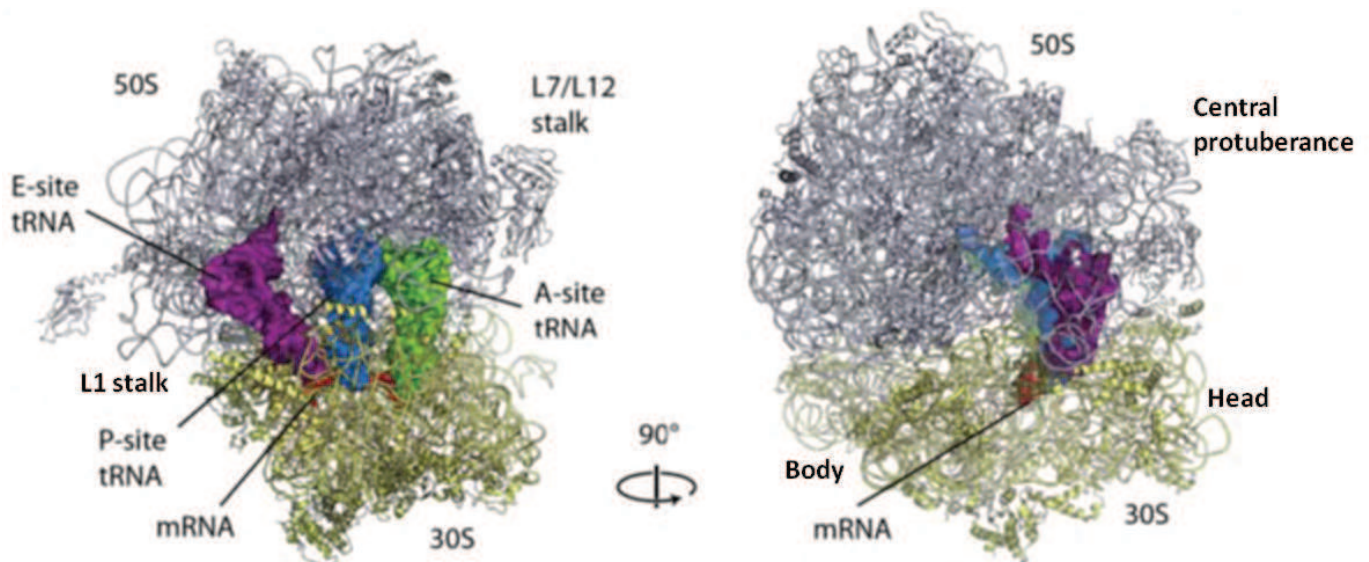
Figure 3: Structural landmarks of the two subunits. The 50S subunit is coloured grey and 30S yellow. The A- P- and E- site tRNAs are coloured in green, blue and yellow respectively. PDB codes: 2WDL, 2WDK.

Using these technological advances, in the first 2 years of the 21st century, several ribosome structures (Ban et al., 2000; Clemons et al., 2001; Nissen et al., 2000; Noller et al., 2001; Yusupov et al., 2001; Yusupova et al., 2001) were published. They revealed that rRNA dominates the functionally important regions such as the PTC, GAC, decoding site for mRNA and the subunit interface. The rRNA is the major workhorse, and is crucial for polypeptide synthesis while the proteins form the scaffold, justifying application of the term ribozyme (RNA based enzyme) to ribosome (Steitz and Moore, 2003).

The two subunits of this ribozyme have characteristic features. 50S has 3 signature points (Fig. 3); central protuberance (CP), the flexible L1 and L7/L12 stalks (Ban et al., 2000) and its rRNA can be divided into seven domains (including the 5S rRNA as domain VII).
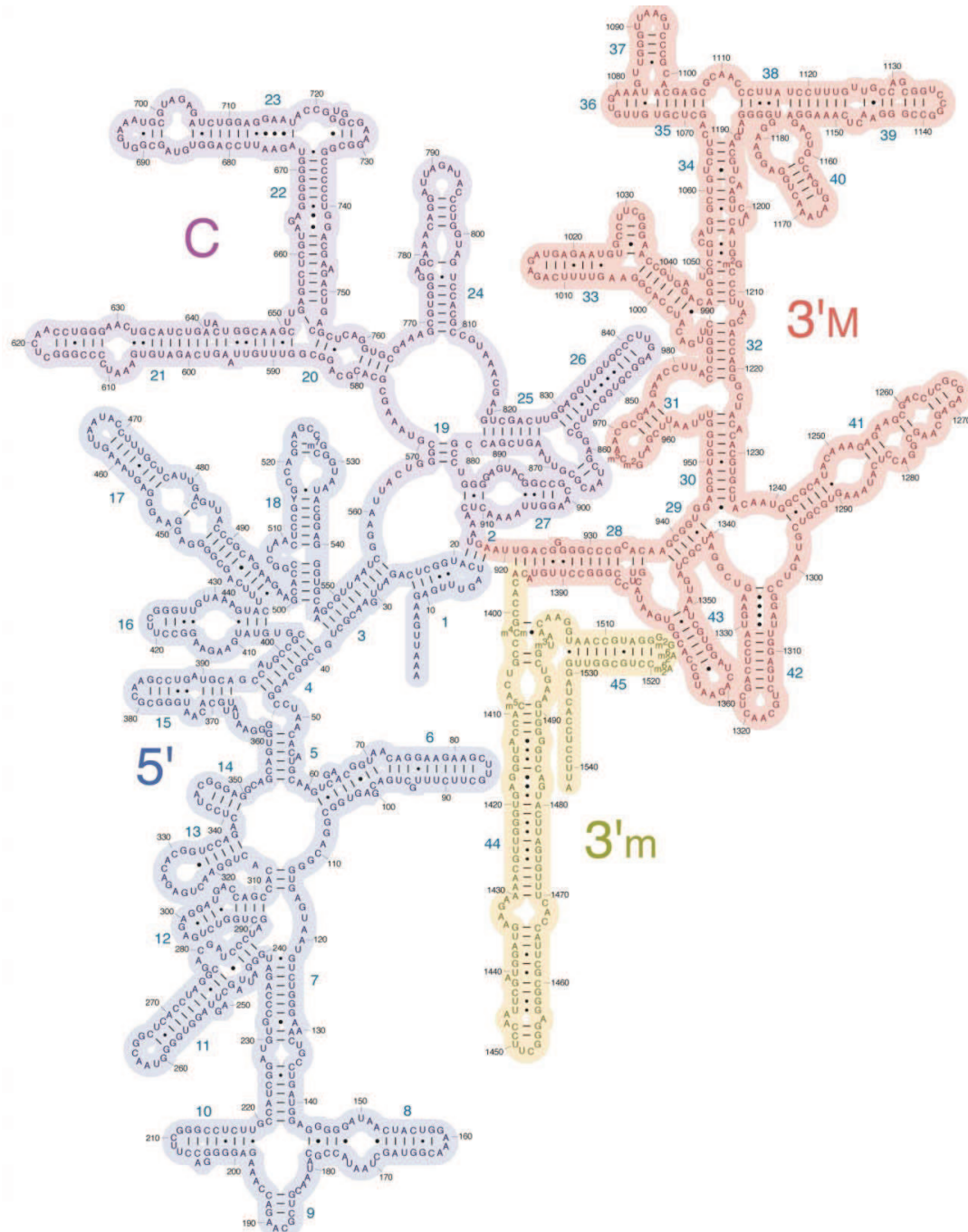
Figure 4: 2D RNA map for the 16S rRNA. The domains are marked in the 16S rRNA.

While the 16S rRNA of the small subunit can be divided into four domains (Fig. 4) and is characterised by the following structural landmarks (Wimberly et al., 2000):

- body (and spur or foot) composed of 5' and 3' minor (h44) domains with proteins rpS4, rpS5, rpS12, rpS16, rpS17 and rpS20;

- the head composed of the 3' major domain and rpS2, rpS3, rpS7, rpS9, rpS10, rpS13, rpS14 and rpS19;

- The platform made by interaction of central domain with rpS1, rpS6, rpS8, rpS11, rpS15 and rpS18.

Both subunits need to act in a synchronised manner for translation. Several interface contact points exist between these subunits called bridges which convey the "message" from one subunit to the other. These bridges are highly dynamic to allow conformational rearrangements to the components of this huge machinery, without breaking the two subunits apart.

1.3.1 Unravelling the ribosome

Crystal and cryo-EM structures of ribosomes have been determined in a variety of states with mRNA, tRNA, factors and antibiotics (Beckmann et al., 2001b; Chandramouli et al., 2008; Gao et al., 2003; Halic et al., 2004; Klaholz et al., 2004a; Klaholz et al., 2003; Nilsson et al., 2004; Simonetti et al., 2008; Spahn et al., 2001; Spahn et al., 2004a; Wilson et al., 2005; Yusupova et al., 2001). They have led to a deeper understanding of this machinery and mechanism of translation, down to the residues involved in every step. Below is a domain-by-domain description of the ribosome catalysis.

1.3.2 Decoding centre

At the interface of the small subunit is the decoding centre where the codon-anticodon interactions take place. The binding of the correct aminoacyl tRNA to A site, as dictated by the mRNA, followed by GTP hydrolysis by EF-Tu, dissociation of the factor and the final movement of the tRNA into PTC is called accommodation and is the rate-limiting step in elongation. There are three universally conserved residues in 16S rRNA, A1492, A1493, G530 (*E.coli* numbering) and protein S12 involved in the recognition of these interactions (Ogle et al., 2001).

In the vacant ribosome A1492 and A1493 appear to be stacked in h44 of the 16S rRNA. It was initially observed that on encountering a cognate tRNA these residues undergo conformational changes. A1492 and A1493 flip out from h44, but retain stacking interaction

with each other while contacting the minor groove of the first two basepairs (codon-anticodon minihelix) while at the third (wobble) position the interactions are weaker, in accord with genetic code degeneracy (Ogle et al., 2001). G530 undergoes a change of conformation from *syn* to *anti*, allowing it to probe minor groove of the second and third base pairs. Residue A1913 of 23S rRNA interacts with the codon–anticodon helix (Selmer et al., 2006).

But lately, new structural insights have been obtained wherein, these conformational changes in the decoding centre with the A1492, A1493 and G530 residues were induced not only upon cognate or near-cognate tRNA binding (Demeshkina et al., 2012). Instead, it was shown that these residues react in an exact similar way to near-cognate tRNA as well.
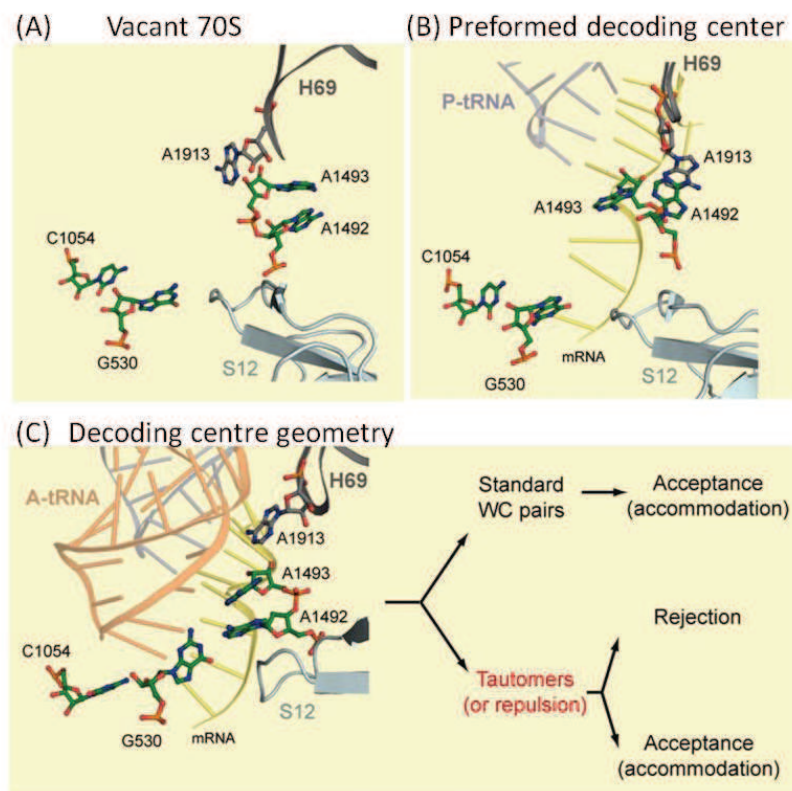
Figure 5: Decoding by the ribosome. (A) and (B) show the conformation of the universally conserved residues at the decoding centre, in the absence and presence of a cognate or near-cognate tRNA. (C) Depicts the formed or closed state of the centre. The discrimination of cognate, near-cognate tRNAs and non-cognate tRNAs is suggested to be based on the energy cost imposed by a rigid decoding centre, stressing on Watson-crick (WC) geometry. Adapted from (Demeshkina et al., 2013).

Moreover, the resulting rigid conformation of decoding centre is enforcing Watson-Crick geometry upon the codon-anticodon duplex. Therefore, in the enhancement of the previous induced-fit hypothesis (Schmeing et al., 2005b), the ribosome undergoes conformational changes upon tRNA binding and these very changes are needed to restrain codon-anticodon duplex geometry to force discrimination between cognate, near-cognate or non-cognate tRNA (Fig. 5) (Demeshkina et al., 2013).

Furthermore, the tRNA at the A-site adopts conformation that features a bend at the anticodon stem, thus facilitating interaction between the decoding centre and the factor binding site where EF-Tu is bound. This allows GTP hydrolysis and release of EF-Tu on correct codon-anticodon recognition. In turn, the tRNA relaxes to its open conformation into the PTC (Schmeing et al., 2009).

The energy of these interactions with cognate as well as near-cognate tRNA induces 30S domain closure, moving the shoulder domain of the 30S towards the neck by ~3Å in 70S ribosome complex with tRNA in A, P and E sites (Jenner et al., 2010). This "domain closure" was also observed in studies with isolated 30S complexes but to a larger extent with rotation of head towards the shoulder of the small subunit. Nevertheless, this rate of "domain closure" tends to dictate the accuracy of the peptide incorporation, as mutations that facilitate this movement have been observed to decrease its accuracy whereas mutations that slow down "domain closure" increase the accuracy (Ogle et al., 2002).

In case, a non-cognate tRNA is delivered to the A-site, it will not undergo accommodation due to its weaker interactions with the decoding centre. This will prevent hydrolysis of GTP associated with its EF-Tu and thus, will result in tRNA release. This proofreading step of accommodation and domain closure contributes to the high fidelity of ribosomal protein synthesis (Wohlgemuth et al., 2010).

1.3.3 The peptidyl transferase centre (PTC)

This centre in the large ribosomal subunit is formed by 23S rRNA residues U2506, G2583, U2584, and U2585 (*E.coli* numbering). No metal ion or ribosomal protein is directly involved in the catalysis of the peptide bond formation. The first requirement to proceed towards peptide bond formation is the proper positioning of the two substrates; the incoming aminoacyl tRNA and the tRNA in the P-site harbouring the growing peptide chain.

With tRNA present in the P-site, the 2'-hydroxyl group of the A76 residue functions as a vital proton shuttle in substrate-assisted catalysis (Weinger et al., 2004). The α-amino group from the incoming charged tRNA in A-site gives the proton to the 2'-hydroxyl group of A76 which in turn, provides a proton to the leaving 3'-hydroxyl group of tRNA on completion of peptide bond formation (Fig. 6). The tetrahedral transition state stabilization is possible by hydrogen from the polar water molecule, coordinated by rRNA bases. The details, however, are subject to debate.

Also, some proteins have been implicated in aiding the process of peptide bond formation. The N-terminal tail of RPL27 is ordered in the PTC where it interacts with the tRNA substrates. And RPL16 interacts with the acceptor arm of A-site tRNA and becomes ordered in structure.

In order to prevent erroneous peptide hydrolysis, the ester linked carbonyl carbon of the peptidyl tRNA is protected from nucleophilic attack via water molecules by U2585, A2451 and C2063 bases of 23S rRNA. Also, to prevent intramolecular transesterification from 3'- to 2'-oxygen, 2'-hydroxyl group of A2451 may be essential (Lang et al., 2008). However, no rRNA residue has been shown to be directly involved in this catalysis.
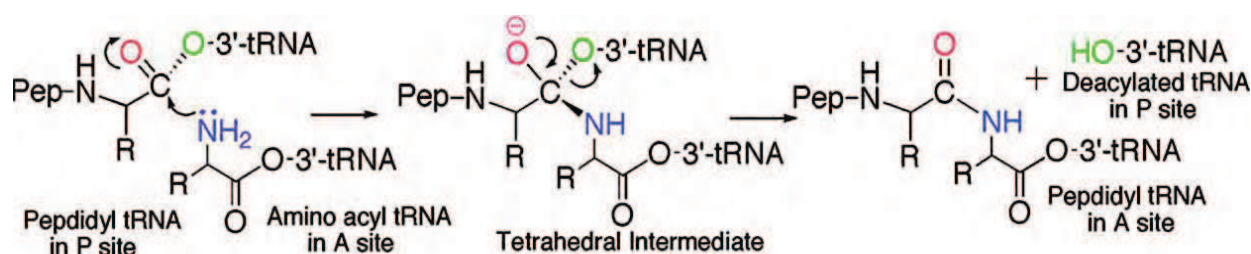
Figure 6: Peptidyl transferase reaction. Nucleophilic attack of the α-amino group of the aminoacyl tRNA at the A-site on the carbonyl carbon of the peptidyl-tRNA at the P-site. Adapted from (Carrasco et al., 2011).

1.3.4 Peptide exit tunnel

The growing peptide chain passes through a 80 Å long and 10–20 Å wide, universally conserved tunnel (Nissen et al., 2000) in the large ribosomal subunit. The tunnel, like other active sites of ribosome, is predominantly composed of rRNA core and some protein components. RPL4, RPL22 along with a bacteria-specific extension of RPL23 and the 23SrRNA segments form the tunnel wall in bacteria. RPL39e replaces the RPL23 near the

tunnel exit in eukaryotes (Harms et al., 2001). RPL4 and RPL22 form a constriction of the tunnel in both 50S and 60S, about 30 Å downstream from the PTC.

Upon emerging on the solvent exposed side, the nascent polypeptide chain interacts with several proteins like chaperones and post-translational modification enzymes (Berndt et al., 2009). In bacteria, a deformylase enzyme binds at the N-terminus of the emerging peptide to cleave off the formyl group bound to the first Met residue. It is recruited to the ribosome by RPL32, present on the exit side of the peptide channel. A trigger factor (a bacterial protein chaperone) is known to interact with RPL23 and RPL29 at the end of the tunnel to assist in proper folding. In eukaryotes, L31e is present instead of RPL32 which interacts with protein Zuotin in yeast (Ben-Shem et al., 2011). Zuotin is a part of eukaryote-specific chaperone complex and is involved in co-translational folding.

Moreover, the ribosome-nascent chain complex, as a whole, can be targeted to protein conducting channel in the membrane of cellular components like mitochondria and the endoplasmic reticulum. This happens only if the protein being synthesised expresses a signal sequence in the nascent peptide chain. This signal sequence is recognised by SRP (Signal recognition particle) thus, directing the targeting of the complex to its destination.

Functionally, the tunnel was initially considered as an inert conduit for the peptide. Lately, it has been shown that the tunnel might be involved in protein folding (Lu and Deutsch, 2005) and translational regulation (Tenson and Ehrenberg, 2002). It has an overall electronegative potential that might allow stalling or reducing translation by interacting with long stretches of positively charged residues, such as arginine or lysine. Eukaryote-specific insertion of RPL4 and RPL17 allow contact with the fungal arginine attenuator peptide (AAP) and the human cytomegalovirus (hCMV) peptide chains in 80S affecting their rate of translation (Bhushan et al., 2010), while in prokaryotes RPL22 mediates translation stalling interactions (Seidelt et al., 2009). This translation pausing could be a method for regulating translation, ensuring efficient membrane targeting of certain proteins and in some cases in prokaryotes, confirming the correct splice variant to be translated (due to coupling of transcription and translation).

1.4 Structural dynamics and coordinated ribosome movements

During translation, this giant ribonucleoprotein complex needs to be flexible and dynamic enough to allow binding of factors, tRNA and mRNA translocation. The biochemical and footprinting experiments, in the late 1980's, first suggested that intersubunit movements allow translocation in two steps. The first step involves movement of tRNA to form a hybrid state where the anticodon stems of the two tRNAs are still in A- and P- sites, while their peptide terminals have moved to the P- and E-sites, depicted as A/P and P/E, respectively. These hybrid intermediate configurations serve in lowering the activation barrier for translocation (Dorner et al., 2006). Subsequently, the second step, catalysed by an elongation factor, allows their movement completely to P-and E-sites, coupled to mRNA movement on the 30S (Moazed and Noller, 1989; Spirin et al., 1987).

The two ribosomal subunits act in a coordinated fashion to allow rotation and swivelling of the small subunit, relative to the large one, thereby ensuring translocation of mRNA and tRNAs. Using cryo-EM studies and solution FRET analysis, the small subunit was observed in a different conformation on binding EF-G (Chen et al., 2013). Frank and Agrawal showed that the small subunit rotates counter-clockwise with respect to the large subunit during translocation, giving rise to the ratchet like mechanism of translocation (Frank and Agrawal, 2000). On ratcheting, the L1 stalk moves by 40 Å, the small subunit body/platform rotates by 5-8° and the head undergoes 12°-15° swivel-like rotation (Fig. 7) (Connell et al., 2007; Spahn et al., 2004a).

In order to allow this intersubunit movement, intersubunit bridges at the extremities are disrupted while those located near the rotation axis (B2a-c, B3, B5, B7a) are essentially maintained (Gao et al., 2003). Later spontaneous fluctuation between this hybrid and classical states was observed for ribosomes containing deacylated tRNA in the P-site, even in the absence of EF-G (Cornish et al., 2008). Binding of the factor and movement of the acceptor stem of deacylated tRNA into E-site seems to stabilise the rotated, hybrid state. This rotation has been seen not only for EF-G, also other GTP binding factors (IF2, EF-G, RF3 and RRF and their eukaryotic orthologs), which makes it a universal mechanism for all the steps of translation (Allen et al., 2005; Gao et al., 2005; Klaholz et al., 2004b).

Intersubunit movement is synchronised with the rotation of the head of the small subunit (Connell et al., 2007; Gao et al., 2003). During the first step of translocation, the head rotates by 12°, in order to prevent breaking apart of the contacts between tRNA and ribosome. Interestingly, an additional movement of the head is predicted for the second step of translocation to create about 20 Å space for anticodon stem loop of tRNA to pass from P- to E-site (Schuwirth et al., 2005). This movement is reversed, on binding of mRNA and tRNA to the P-site, leading to classical pre-translocation state (Berk et al., 2006).
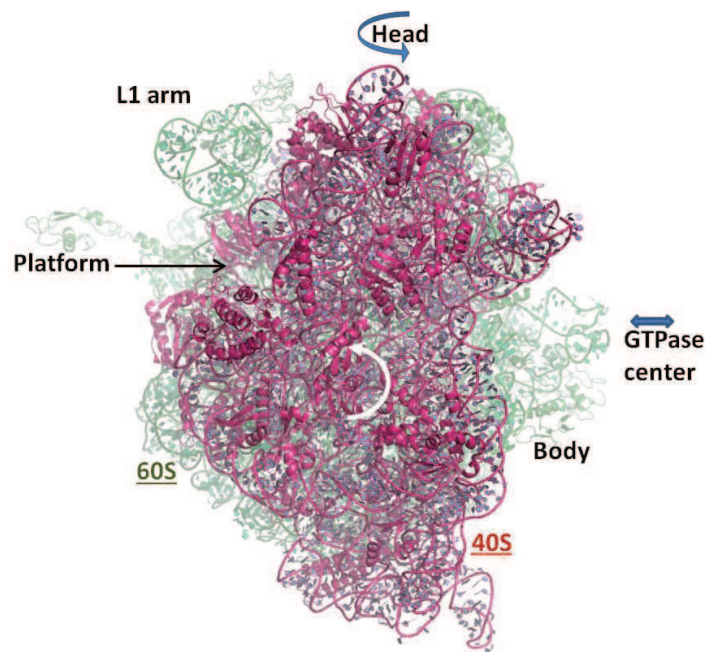
Figure 7: Intersubunit movements. The arrows indicate major conformational rearrangements in ribosome including rotation of 30S body and swivelling of the head. On the 50S the L1 proteins form the dynamic cluster. 50S and 30S PDB codes: 2WDL, 2WDK.

On the large subunit, L1 and L11 stalks are the most flexible regions. L1 stalk comprises RPL1, helices H76, H77 and H78 of 23S rRNA and has been observed in "closed" and an "open" conformation. The L1 stalk moves 30-40 Å inwards as compared to an open conformation in the presence of E/E site tRNA (classical state). In the P/E hybrid state it moves 15-20 Å further closer to the small subunit (closed) conformation. To allow the release of the deacylated tRNA from the E site, it moves to an open conformation. L11 stalk is associated with the GAC, on the opposite side of the 50S from L1 stalk. It is composed of helices 42, 43 and 44 of 23S rRNA and protein L11. L11 has been observed in an "inward" and an "outward" position (Stark et al., 1997; Stark et al., 2000) in ribosome structures with tRNA.

Apart from these global and easily noticeable movements, there exist local rearrangements like in the PTC and the decoding centre, making ribosome a highly adaptable entity. Moreover, these conformational changes during translation and the various translational active sites composing PTC, decoding centre, sites for tRNA binding, GAC and the peptide tunnel are similar and conserved in ribosomes from all domains of life as elaborated with eukaryotic ribosome structure in the next section.

1.5 Eukaryotic ribosome and its biogenesis

The eukaryotic ribosomes are considerably larger as compared to their prokaryotic counterparts (Table 2), probably representing a higher level of regulation due to compartmentalization and increasing complexity at the cellular level.

Prokaryotes		Eukaryotes (Mammalian)	
2.3 MDa 54 proteins and 3 rRNA		4.3 MDa in humans 80 proteins and 4 rRNA	
50S	**40S**	**60S**	**40S**
33 proteins 23S rRNA 5S rRNA	21 proteins 16S rRNA	47 proteins 28S rRNA 5.8S rRNA 5S rRNA	33 proteins 18S rRNA

Table 2: Composition of prokaryotic and mammalian ribosomes.

Their synthesis is a multi-step, error-prone process. It requires coordinated activity of more than 200 non ribosomal *trans*-acting factors (Warner, 1999). Ribosome biogenesis starts in the nucleolus where the rRNAs components are synthesised. Except for 5S rRNA, the rest of the three rRNA components are synthesised by RNA Pol I from a single transcription unit. The fourth rRNA (5S) requires RNA Pol III. Most of these pre-rRNA transcripts are cleaved co-transcriptionally and loaded with a set of proteins, generating pre-40S and pre-60S ribosomes (Kos and Tollervey, 2010).

However, in some cases, a full-length transcript (35S) is generated and assembled into a 90S pre-ribosome. This 90S particle contains no large subunit proteins and upon cleavage of its rRNA component, it releases the pre-40S particle that follows a maturation pathway independent of the large subunit assembly (Fig. 8) (Grandi et al., 2002). Various *trans*-acting factors assist the functionally inactive pre-60S and pre-40S to exit the nucleus via the nuclear pore (Nissan et al., 2002). In exponentially growing yeast cells, about 30 of these pre-ribosomes are exported every second, depending upon nutrient availability (Warner, 1999). Also, during maturation, the crucial aspect is the nucleo-cytoplasmic partitioning of the events. For example, the acquisition of 40S beak most likely occurs once pre-40S has reached the cytoplasm (Schafer et al., 2006). Finally, as observed for most

cellular pathways, several GTPases and AAA-type ATPases ensure recycling of the factors involved in late maturation steps (Ulbrich et al., 2009). Nevertheless, the identity and function of many participating factors (unclassified factors in Fig. 8) still remains to be determined.
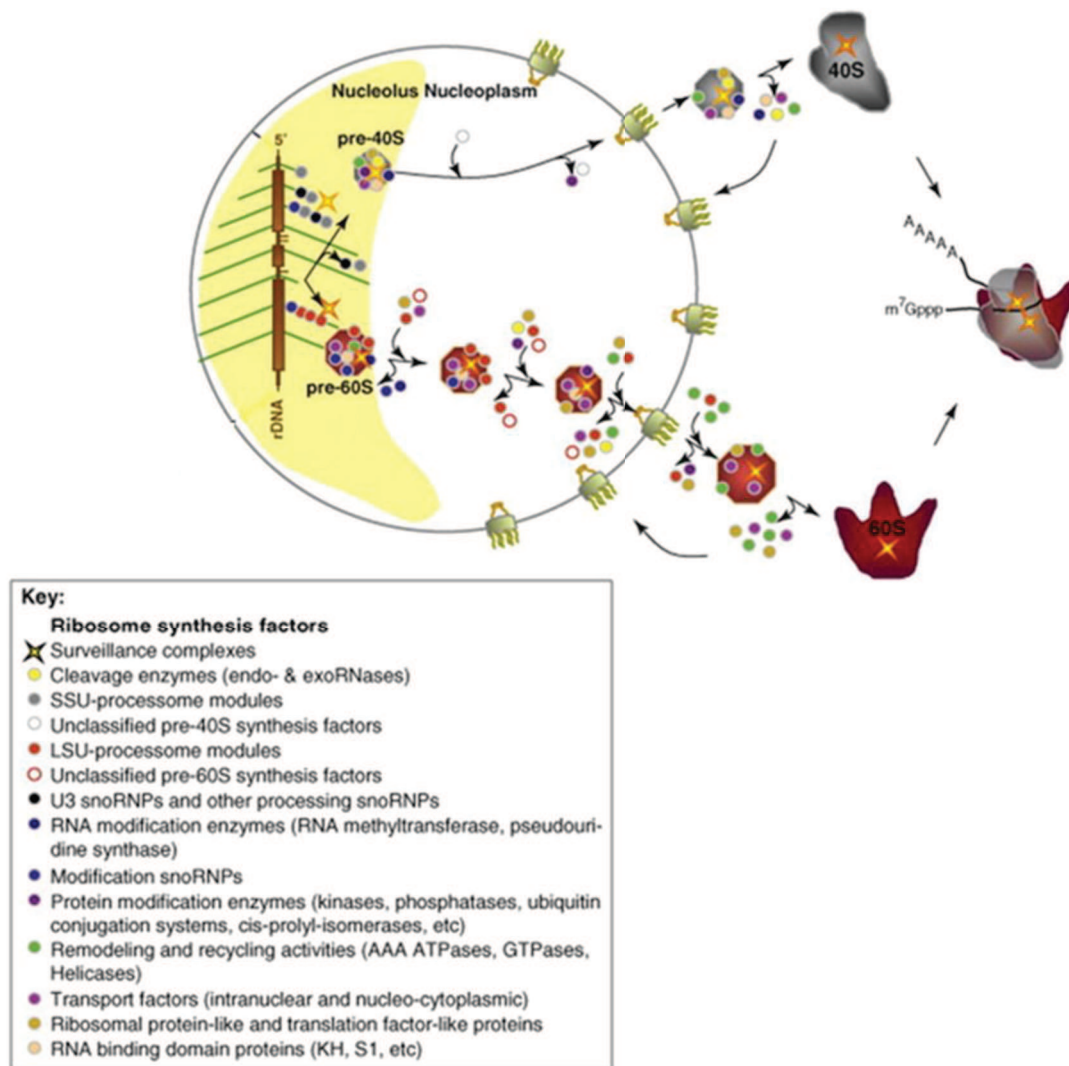
Figure 8: Ribosome biogenesis. The actively transcribed rDNA unit is depicted as a reminiscent of its visualization by Miller chromatin spread (Scheer et al., 1997), with the "trunk" (brown) representing the rDNA locus and the "branches" (green) corresponding to nascent pre-rRNA transcripts. The pre-40S and pre-60S entities exit the nucleus via the nuclear pore (marked in green) and the rest of the maturation takes place in the cytoplasm. Several ribosome synthesis factors are colour coded as indicated in the key. Adapted from (Lafontaine, 2010) with some changes.

1.6 Eukaryotic ribosome structure

Despite several ground-breaking studies on ribosome crystal structures, the atomic structure of eukaryotic ribosome has been an enigma until a few years ago. At the time we started this project, only the yeast ribosome crystal structure was known at 4.15 Å resolution (Ben-Shem et al., 2010). However, single-particle cryo-EM had been the favourable approach for the higher eukaryotic structures. This could be due to the inherent flexibility of these ribosomes which is a hindrance for crystallization. Nevertheless, these studies for canine (Chandramouli et al., 2008), human (Spahn et al., 2004b), wheat germ (Becker et al., 2009) ribosomes have helped in localization of eukaryote-specific proteins like RACK1, S19e on the SSU and L30e on the LSU; and some part of the rRNA additional nucleotide clusters called as expansion segments (ES).

The landmark for these studies was achieved in the past 3 years.

Crystal structures:

- The yeast ribosome structure at 3 Å (Ben-Shem et al., 2011)

- Structures of the large and small ribosomal subunit of *Tetrahymena thermophila* at 3.5 Å and 3.9 Å respectively (Klinge et al., 2011; Rabl et al., 2011).

- Rabbit 40S complex at 11 Å (Lomakin and Steitz, 2013).

Single particle cryo-EM structures:

- Wheat ribosome structure at 5.5 Å (Armache et al., 2010)

- *Trypanosoma brucei* ribosome structure at 5 Å (Hashem et al., 2013)

- Human and *Drosophila melanogaster* ribosome structures at 5.4 Å (Anger et al., 2013)

- Yeast ribosome bound to CrPV IRES structure at 4 Å (Fernandez et al., 2014)

- Yeast mitoribosomal large subunit at 3.2 Å (Amunts et al., 2014)

- Pig ribosome structure bound to Sec61 complex at 3.4 Å (Voorhees et al., 2014)

Technically, it is interesting to note that so far, high-resolution crystal structures have been reported only for single-celled eukaryotic ribosomes. For mammalian ribosome structural studies, obtaining a highly homogenous, crystallisable sample still remains a major hindrance. In contrast, the leap from sub nanometre resolution to attaining close to atomic resolution in cryo EM studies shows how the two techniques complement each other. Even with inherent plasticity in the sample, the latter allows obtaining detailed structural information of the sample. This advancement in resolution has been made possible due to new softwares and highly stable equipments for data collection nowadays (detailed in methods section).
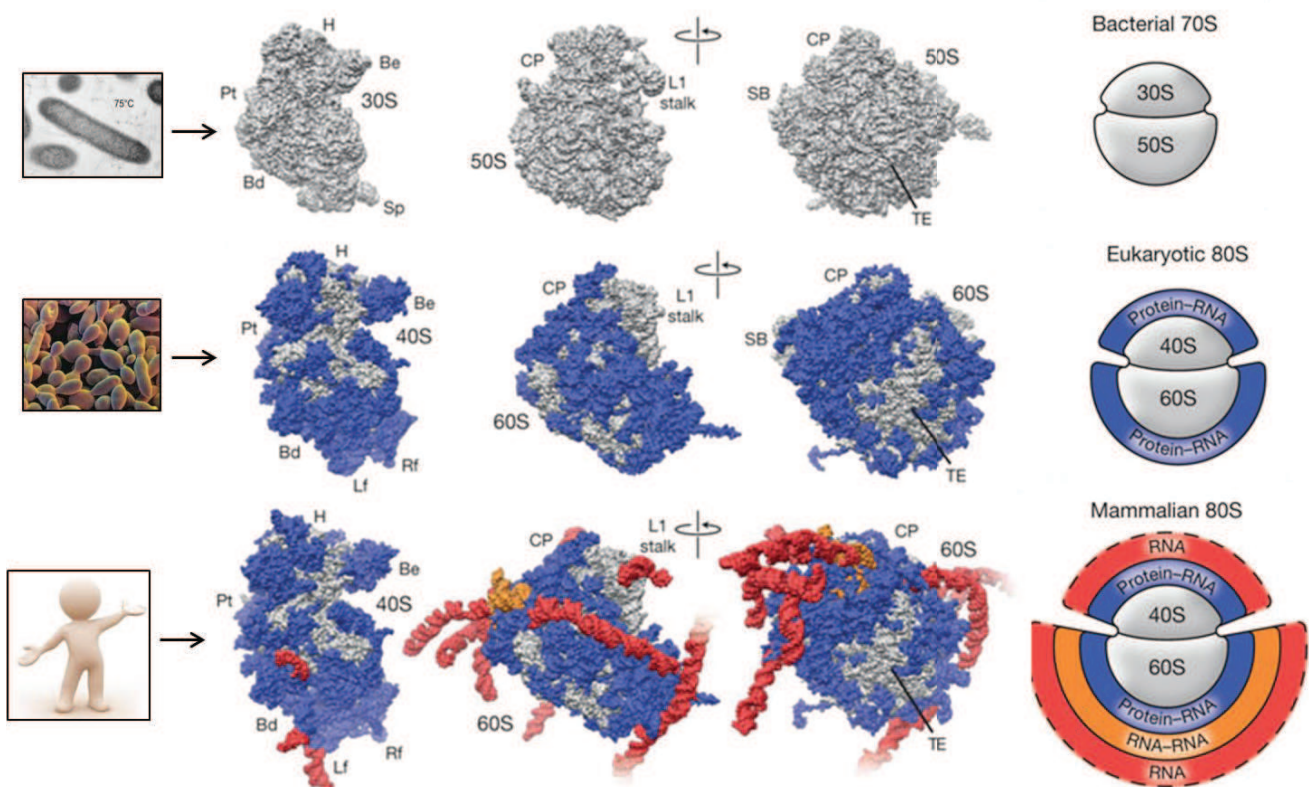
Figure 9: Conservation of the ribosome core in different domains of life. In grey is the bacterial ribosome core, in blue the yeast specific protein and RNA addition, in orange the RNA-RNA layer consisting of rigid RNA moieties specific for mammals and in red the highly flexible RNA expansion segments specific for mammalian ribosomes. Adapted from (Anger et al., 2013).

These structures emphasise the conservation of ribosome core in all domains of life. On comparison of yeast, human and bacterial ribosomes, it was established that the central part of the ribosome possesses a conserved protein-RNA structure (Anger et al., 2013). The

species-specific addition of RNA and proteins are majorly on the solvent exposed sides (Fig. 9), as explained in the next paragraphs.

1.6.1 Eukaryotic ribosomal proteins

The eukaryotic ribosome has 13 eukaryotic specific proteins, 35 proteins with bacterial homologues and 32 with only archaeal homologues. Moreover, most of the conserved proteins have eukaryote-specific extensions extending from the globular core of the protein, which aid in long-distance interactions (50–100Å) as observed for S5, L4, L7, and L30 proteins. The interesting aspect is the intertwined interaction of expansion segments (ES) with ribosomal proteins which do not exist in prokaryotes. Instead, they have simple tertiary protein-RNA interactions mediated by positively charged extensions of ribosomal proteins. These ES-protein interactions tend to stabilise the RNA segments, like for ES7L interacting with RPL28e.

In addition, some proteins form a network of interactions with their eukaryote-specific extensions, like RPL21e. It forms secondary structure elements (intermolecular shared β-sheets) with RPL30, RPL7 and RPL18A; and it anchors ES12 on the surface of 60S subunit thus, sandwiching ES12 along with eukaryote-specific protein RPL29. RPL18A also stabilises ES39 and its associated proteins (Klinge et al., 2011). Also, it forms a cradle with RPL14 oriented towards ES39 and a eukaryotic specific extension of RPL13A. RPL6, positioned on the top of ES39 contacts ES7 and ES39m and forms inter-protein shared β-sheets with RPL14e (Ben-Shem et al., 2011).

Structurally, several proteins have conserved domains involved in keeping this huge protein-RNA moiety together. Proteins like RPL6, RPL13A and RPL14 possess SH3 (Src homology 3) domains characterised by a β-barrel fold that consists of five or six β-strands arranged as two tightly packed anti-parallel β sheets. These SH3 domains are used for both protein-protein as well as protein-RNA interactions forming eukaryote-specific clusters. Furthermore, four zinc-finger (ZnF) proteins, rpS26e, rpS27e, rpS29e and rpS31e, have been identified in the small subunit. rpS29e is a close homologue of spS14p (bacterial) and is embedded in the core while the other three proteins have archaeal homologues (Lecompte et al., 2002). Another protein of the small subunit, rpS26e displays a unique fold related to

Zn-F domains termed as FYVE domain (characterized by two small β hairpins, followed by an α-helix). This domain might be involved in 3'-end processing of 18S rRNA (Rabl et al., 2011).

Quite a few of the ribosomal proteins participate actively in ensuring proper biogenesis. rpS27a, located at the beak of the 40S, is expressed as a ubiquitin-fusion protein (Lacombe et al., 2009). The cleavage of ubiquitin is a prerequisite for proper ribosome function, in the absence of which it would prevent tRNA binding due to steric hindrance. RPL40 similarly possesses a ubiquitin domain at its N-terminus (Lacombe et al., 2009). The α-sarcin-ricin loop (SRL) and elongation factor binding site are present in close proximity of RPL40. In the absence of ubiquitin cleavage it would sterically block this site. This regulation mechanism averts immature ribosome assembly (Klinge et al., 2011).

Some eukaryotic proteins are similar to the prokaryotic proteins, like rpS1e which substitutes for bacterial rpS6p and rpS18p proteins. It consists of an N-terminal β-barrel domain related to rpL33p and a C-terminal, α/β domain resembling rpS6p. However, rpS1e is shifted in position along the RNA as compared to its position in prokaryotic ribosomes, and is in contact with ES7 stabilizing its conformation (Rabl et al., 2011). Another protein with conserved structure is rpS6e. It possesses a β-barrel domain at its N-terminus, which is closely related to bacterial rpL25p. It wraps around the rRNA and has a phosphorylated C-terminal extension in most eukaryotes.

Furthermore, rpS7e, resembles the fold of NusA (bacterial transcription regulator), consisting of two evolutionarily conserved K homology domains (KH). The KH domain binds RNA and can dictate RNA recognition. Here, it inserts into eukaryotic specific rRNA segments (ES6A, ES6B, ES6E), thus binding to the 18S rRNA (Fig. 10). Also, it forms an extended β-sheet with rpS22e (like some proteins in the large ribosomal subunit). Another eukaryote-specific interaction is exemplified by rpS4e. It consists of three domains and an N-terminal extension which is buried deeply in the rRNA (Rabl et al., 2011).

The small subunit head is characterised by the presence of RACK1 (Receptor for activated C kinase 1) on the solvent exposed side. It has a seven-bladed β-propeller structure with six-fold pseudo-symmetry. It is a scaffold protein which is involved in connecting signalling transduction pathways with translation by recruiting several proteins. It communicates with almost all signalling pathways in the cell including protein kinase C,

cAMP/PKA pathway, receptor tyrosine kinases, transmembrane receptors and the Src family of non-receptor protein tyrosine kinases. On the small subunit, it is known to interact with acetylcholinesterase, 14-3-3 protein, androgen receptor and many others.

There exists a deep structure-function correlation in the ribosome (Fig. 10) like at the mRNA exit and entry sites. In order to permit mRNA scanning for initiation, helix 16 (h16) of the small subunit needs to be flexible as opposed to its rigid bent conformation in prokaryotes. An extension in protein S4 in prokaryotes interacts with h16 while in eukaryotes, this domain of S4 is absent. Moreover, proteins present at this site: S30e, S3 and S5; assist in mRNA-scanning.

Functionally, only limited information is available for ribosomal proteins. So far, the best-studied are a set of clinical phenotypes arising from ribosomopathies. They are genetic abnormalities in eukaryotic ribosomal proteins synthesis, leading to impaired ribosome biogenesis and function. Diamond Blackfan Anemia is caused by mutations or deletions in one or several ribosomal proteins, including rpS7e, rpS10e, rpS17e, rpS26e, rpS19e, rpS24e and rpS27a. It is characterised by anaemia, macrocytosis, bone marrow failure (Narla and Ebert, 2011). Furthermore, 5q syndrome is triggered by deletion of one allele of rpS14e, leading to insufficient rpS14e expression. It leads to impaired erythropoiesis and macrocytic anaemia (Narla and Ebert, 2010). Schwachman-Diamond Syndrome is another autosomal recessive disease caused by mutations in the highly conserved SBDS gene which plays a role in ribosome biogenesis and RNA processing. Uncoupling of GTP-hydrolysis from eIF6 also has been speculated to be the cause of this disease (Finch et al., 2011).

These ribosomopathies have also been associated with higher susceptibility to cancer (Narla and Ebert, 2010). In spite of knowing the genotype-phenotype correlation for these diseases, the molecular pathways are not known exactly. It has been postulated that these defects could arise due to the regulation function of these ribosomal proteins or their role in ribosome assembly. Further research needs to be done on structure-function correlation of ribosome components, in order to characterise these pathways.
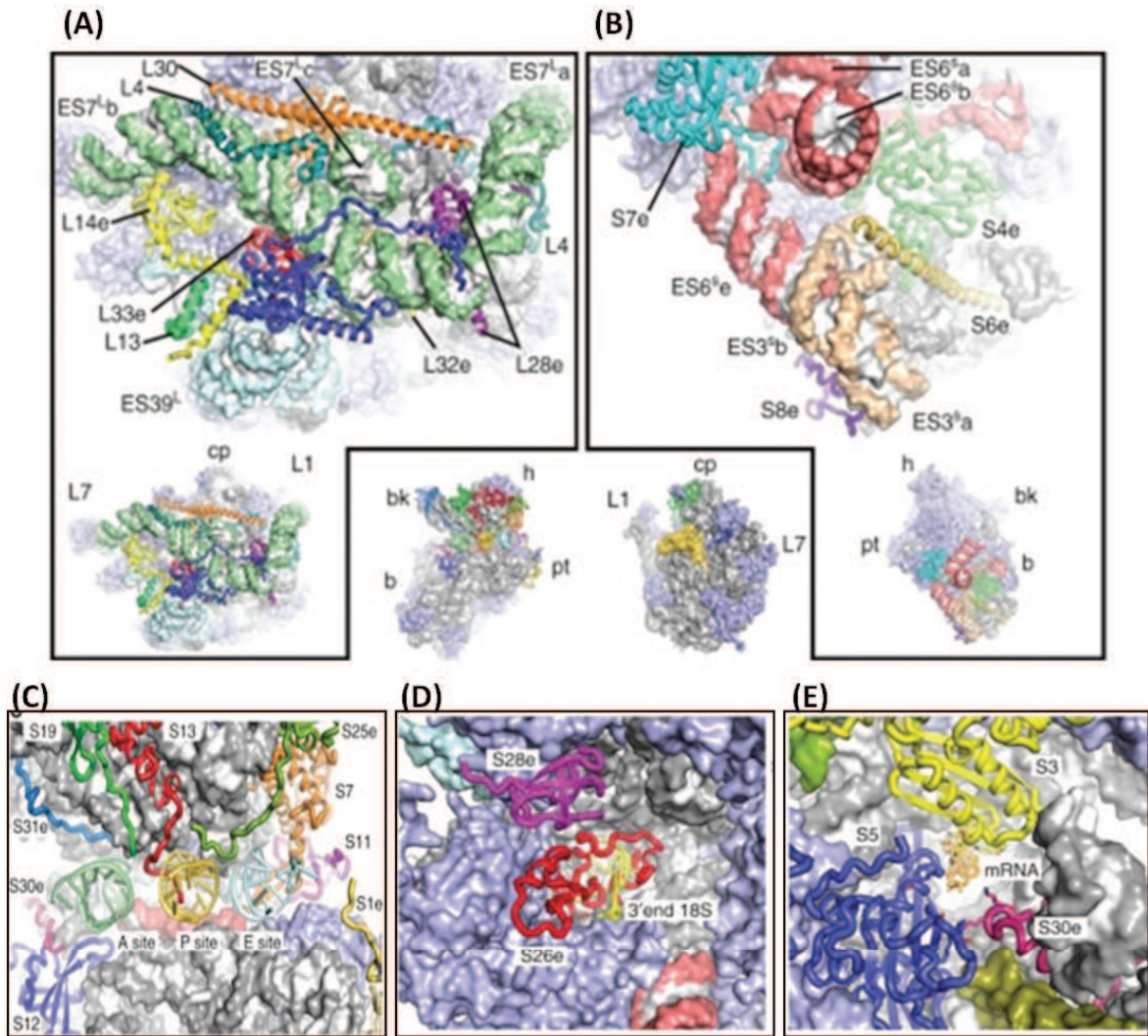
Figure 10: Structure-function correlation in 80S. (A) Interweaving of rRNA and proteins at ES7L and ES39L on the large subunit. (B) ES3S and ES6S on the small subunit with associated proteins. Protein extensions: (C) into the tRNA binding site of 40S. mRNA (D) exit and (E) entry site on the small subunit. Adapted from (Wilson and Doudna Cate, 2012).

1.6.2 Eukaryotic ribosomal RNA

The eukaryotic ribosome has five additional ES (ES3S, ES6S, ES7S, ES9S, and ES12S) and five variable regions (VRs) (h6, h16, h17, h33, and h41) on the 40S and 16 expansion segments (ES3L, ES4L, ES5L, ES7L, ES9L, ES10L, ES12L, ES15L, ES19L, ES20L, ES24L, ES26L, ES27L, ES31L, ES39L, and ES41L) and two VRs (H16–18 and H38) on the 60S (Gerbi et al., 1996). The VRs have diverse sequences amongst the eukaryotes while the ES are addition of RNA sequence as compared to bacterial RNA sequence. Of these, ES7L, ES15L, ES27L, and ES39L are significantly longer in humans (about 100-600 nucleotides longer) as compared to yeast (Cannone et al., 2002).

These expansion segments can be divided into two broad categories. The first being long ES helices protruding from the ribosome like antlers and attached to the ribosome only at their bases. They are highly flexible since almost no density for these long segments has been observed in the cryo-EM reconstructions (Anger et al., 2013). They can adopt different conformations like as observed for ES27L and ES6S. ES27L has been postulated to exist in two conformations (in and out), which might have some functional significance. It has been suggested to play a role in recruiting non ribosomal proteins, like chaperones and modifying enzymes, to the peptide tunnel exit (Beckmann et al., 2001a). ES27L has also been determined to be crucial for cell survival (Sweeney et al., 1994). Due to their accessibility on the surface, it may be synonymous with increased regulation like during initiation or termination (Sonenberg and Hinnebusch, 2009).

The second category is composed of those segments which are tightly associated with ribosomal proteins or other rRNA expansion segments forming eukaryote-specific clusters. One of these clusters is made by ES7L, ES39L, five eukaryotic r-proteins (L6e, L14e, L28e, L32e, and L33e), as well as eukaryote-specific extensions of conserved r-proteins (L4, L13, and L30) present around the solvent-exposed back of the 60S (Fig. 10). ES7La is stabilised by the presence of L28e in wheat germ and *Tetrahymena*, whereas in yeast, this helix is more flexible in due to absence of L28e. Another major ES cluster encompasses ES19L, ES20L, ES26L, and ES31L, associated with L27e, L30e, L34e, L43e (Ben-Shem et al., 2011). ES39L and ES31L form the core of these clusters with stretches of single-stranded rRNA surrounded by ribosomal proteins (Melnikov et al., 2012). The exact role of all these ES has been and remains a long-standing question. Especially, in case of human ribosomes these ES alone increase the molecular weight by one MDa as compared to ribosomes of single-celled eukaryotes. This additional mass must hold some significant role in regulating translation which needs to be explored.

1.6.3 The intersubunit interface

The bridges between the two subunits play a crucial role in the process of translation. There are seven bridges in the core of the ribosome and the subunit interface is highly conserved with very few ES and eukaryote-specific proteins (Ben-Shem et al., 2011). However, in recent studies several unusual bridges have been determined. These bridges are formed by proteins extending from the large subunit and binding to the components of

the small subunit. Like in bacteria one of these bridges is formed by protein L31, located on the CP of the 50S while its C terminal domain is bound to the head of the 30S (Jenner et al., 2010). L19e and L24e in eukaryotes form similar bridges (Ben-Shem et al., 2011).

The eukaryote-specific bridges are concentrated majorly at the periphery of the ribosome and contribute in doubling up the interaction interface. They are predominantly formed by eukaryotic-specific elements. For example, the bridge eB11 formed by protein S8e and ES41L, allows communication between the small subunit platform, just below the mRNA exit tunnel and the large subunit. Consequently, these bridges serve as a link between the functionally important domains.

This is also observed for eB8 composed of ES31L and S1e. ES31L is an essential component of a cluster formed by eukaryotic elements at the back of L1 stalk. This cluster harbours several ES and eukaryote-specific proteins, including extension of protein L23 that forms the universal docking site for factors involved in co-translational regulation. Also, ES31L is attached to the highly mobile L1 stalk that is involved in evacuating exit site tRNA. On the small subunit side, S1e interacts with the components of the mRNA exit tunnel, namely S11 and S26e, thereby linking mRNA-tRNA translocation.

Exceptionally, L41e forms the eB14 bridge, the only eukaryote-specific bridge present in the centre of the ribosome. It seems to be entirely associated with 18S rRNA. However, upon dissociation of the subunits, L41e remains as a part of the large subunit (Melnikov et al., 2012). The binding pocket for L41e is highly conserved in eukaryotes and bacteria, but interestingly, no corresponding protein has been determined in prokaryotes (Schluenzen et al., 2000; Wimberly et al., 2000). The intriguing aspect is the central role of proteins in these eukaryote-specific bridges as opposed to the bridges observed in prokaryotes.

Thus, the prokaryotic and eukaryotic ribosomes have evolved with similar structure-function correlation, but eukaryotic ribosomes need to be investigated in more detail. The function with respect to location of eukaryotic ribosomal proteins, the role and structure of ES are some of the questions that can be addressed only with ribosome and ribosomal complex structures attaining near-atomic resolution.

1.7 Translation Termination

The length of the protein synthesised by the ribosome is dictated by the step of termination. This requires binding of release factors which are divided into two categories:

Class I release factors recognise the nonsense or stop codons, UAA, UGA and UAG (Brenner et al., 1967; Brenner et al., 1965) in the A site. In bacteria RF1 (Release factor 1) recognises UAA and UAG stop codon and RF2 (Release factor 2) recognises UAA and UGA. The two proteins are homologous in sequence and share similar 3D structures (Nakamura et al., 1995; Vestergaard et al., 2001). Thus, there is a mixed specificity with two RFs recognising three stop codons. While in eukaryotes an omnipotent release factor, eRF1 (eukaryotic RF1) commands termination at all three stop codons.

Class II release factors are GTPases and involve RF3 and eRF3. They assist class I release factors and act in a GTP dependent manner. RF3 is known to promote dissociation of class I release factors from the ribosome upon peptide release (Freistroffer et al., 1997). But the role of its eukaryotic counterpart still remains an enigma. It might be required to stimulate the efficiency of eRF1 and promote its recycling as discussed below (Mitkevich et al., 2006).

1.8 RF1 and RF2 structures

The crystal structures of RF2 and RF1 were determined by Vestergaard et al., in 2001 and Shin et al., in 2004. They provided initial structure-function insights for termination. Both factors are composed of 4 domains, with each domain harbouring a functional peptide motif.

- N-terminal domain 1 consists of four α-helices and a 3_{10}-helix. A coiled coil motif is formed by helices α1, α3 and α4. This domain is projected outwards from the main body but hydrophobic interactions, salt bridges and hydrogen bonding keep it secured to the rest of the protein.
- Five stranded anti-parallel β-sheet and two α-helices compose the domain 2. The SPF motif in RF2 and PxT motif in RF1, required for stop codon recognition, is present in a loop between β4 and β5 strands.

- Domain 3, comprising of β6, β7, β8 strands and a long α helix, is situated above domain 2. It harbours the GGQ motif, required for peptidyl tRNA hydrolysis.
- The C-terminal domain 4 connects closely to domain 2 and together with it forms the decoding domain. This super-domain is stabilised by a hydrogen bonded network in the centre composed of conserved polar residues.
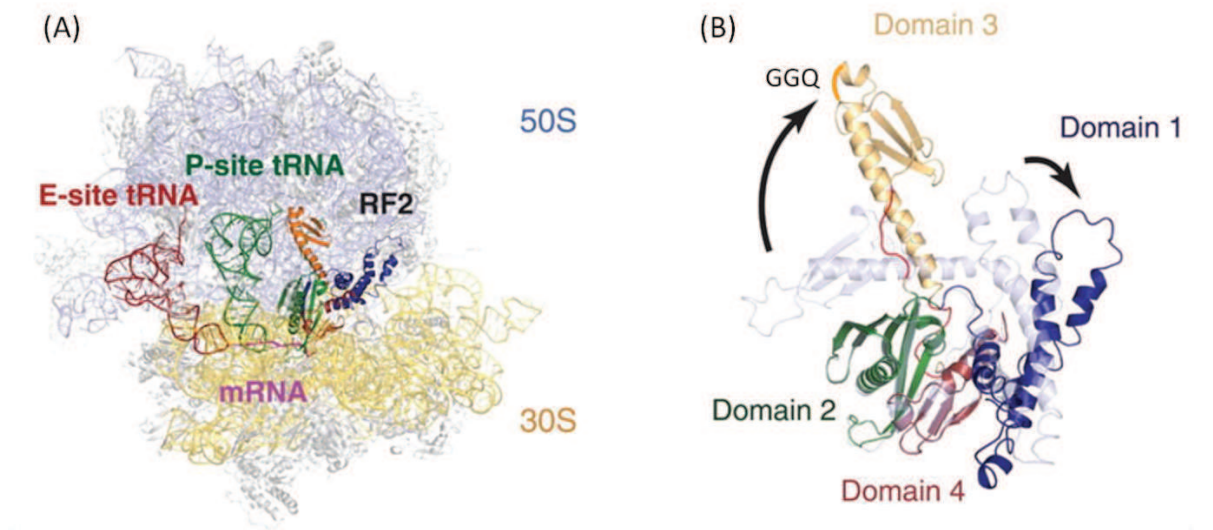
Figure 11: (A) Structure of RF2 bound to the ribosome. RF2 coloured domain wise like in (B). (B) RF2 undergoes conformational change on binding the ribosome. In the background, in light blue is the crystal structure of RF2. The domain 1 (dark blue) and domain 3 (peach with GGQ in orange) movements are marked by arrows. Adapted from (Weixlbaumer et al., 2008).

However, the crystal structures of these factors are highly compact with the two active sites placed only 23Å apart (Fig. 11). The minimum distance to allow interaction between the PTC and decoding site on ribosome is 70 Å. Single particle cryo-EM (Klaholz et al., 2003; Rawat et al., 2003) and SAXS studies (Vestergaard et al., 2005) revealed the existence of an open conformation of RF1 and RF2. The importance of these conformational changes was further verified by crystal structures of termination complexes on ribosome and is explained below.

1.9 Stop codon recognition

It has been suggested that modern life forms originated from RNA based molecular ancestors (Crick, 1968), where RNA served the dual purpose of storing genetic information along with catalyzing biochemical reactions (Kruger et al., 1982; Pace and Marsh, 1985). This gave rise to the concept of RNA world. And the discovery of ribosomes as ribozymes

reinforced this hypothesis. Interestingly, recognition of mRNA sequence by a protein instead of another RNA sequence, is an example how nature has evolved the RNA world into this complex, intricate RNA-protein machinery. Mutational and photo cross-linking studies gave the first insights into the residues involved in stop codon recognition (Brown and Tate, 1994). The PxT (PAT or PVT) and SPF motifs were suggested to act as "anticodons" (Ito et al., 2000; Scarlett et al., 2003). However, it was the structural studies which gave insights into the mechanism of stop codon recognition (Klaholz et al., 2003; Korostelev et al., 2010; Laurberg et al., 2008; Petry et al., 2005; Weixlbaumer et al., 2008).

The residues involved in this recognition are composed of the N-terminal end of helix α5 and the conserved recognition loop [RF1 181-195 (185-199), RF2 201-217 (200-216) in *T. thermophilus (E. coli)*, respectively] between β4 and β5 strands of domain 2. They together form the "reading head". The three nucleotides of the stop codon interact with three elements of the reading head, but not three amino acids, as was predicted by the "tripeptide decoding". (Ito et al., 2000; Nakamura and Ito, 2002) Hence, it is not a direct codon-anticodon reading mechanism between nucleotides and amino acids.
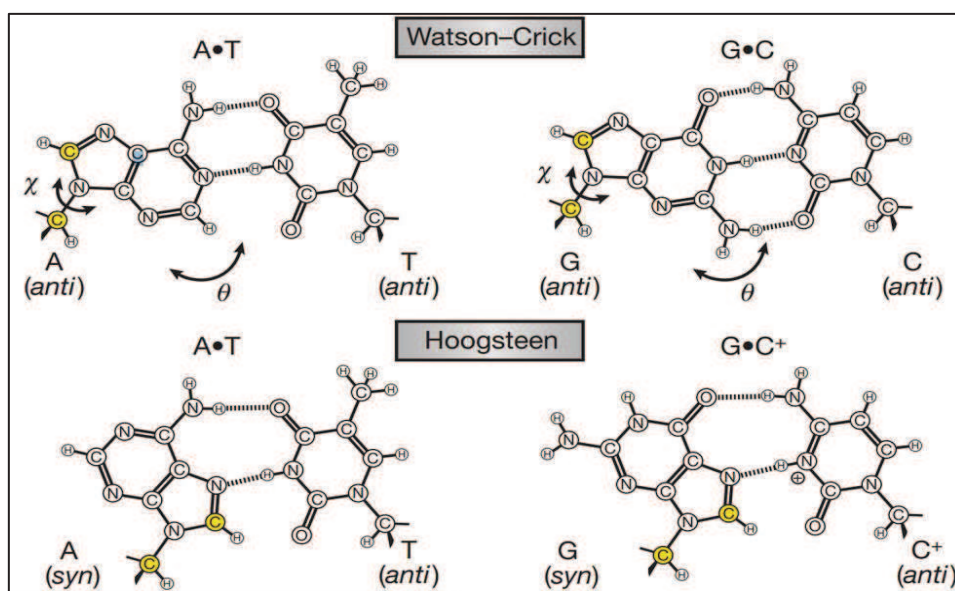
Figure 12: Watson-crick and Hoogsteen interactions between nucleotides. On purine rotation around the glycosidic bond (χ) and base flipping (θ), Hoogsteen geometry can be obtained, affecting C8 and C1 residues, indicated in yellow.

54

First base recognition in stop codon (U1)

U1 is present universally in all three stop codons and thus, the recognition mechanism for this nucleotide is highly conserved. Specific hydrogen-bonds are formed between the N-terminal tip of helix α5 and U1 (Fig. 13). This hydrogen bonding pattern resembles that of the canonical A: U base pair (Laurberg et al., 2008).
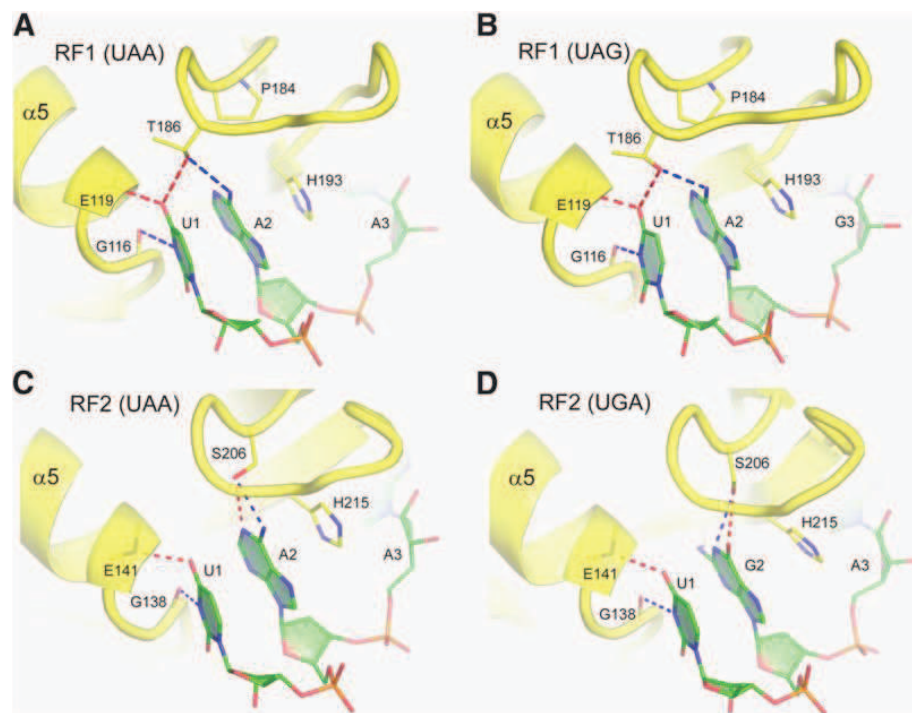
Figure 13: Interactions of the first 2 nucleotides of the stop codons with release factor. Adapted from (Korostelev, 2011).

Second base recognition in stop codon (A2/G2)

Conserved amino acids of the recognition loop dictate the specificity of the second nucleotide recognition. The side-chains of Thr of the PxT motif and Ser of SPF motif in RF1 and RF2, respectively form hydrogen bonds with the A or G in second position (Korostelev et al., 2010; Laurberg et al., 2008). Thr interacts with U1 and the second residue (Fig. 13). However, Ser does not interact with U1 (Weixlbaumer et al., 2008). Apart from this, other residues of the conserved loop are involved in the stop codon specificity in this region. This was demonstrated by simple swapping of the recognition loop of RF1 with that from RF2, rendering RF1 nonspecific for A or G at the second position (Young et al., 2010). Interestingly, it was determined that apart from interacting with Watson-Crick edges, these

factors also interact with Hoogsteen edges (Fig. 12) of codon nucleotides. Glu187 and Asp209 might interact with the Hoogsteen edge of the second nucleotide (Young et al., 2010).
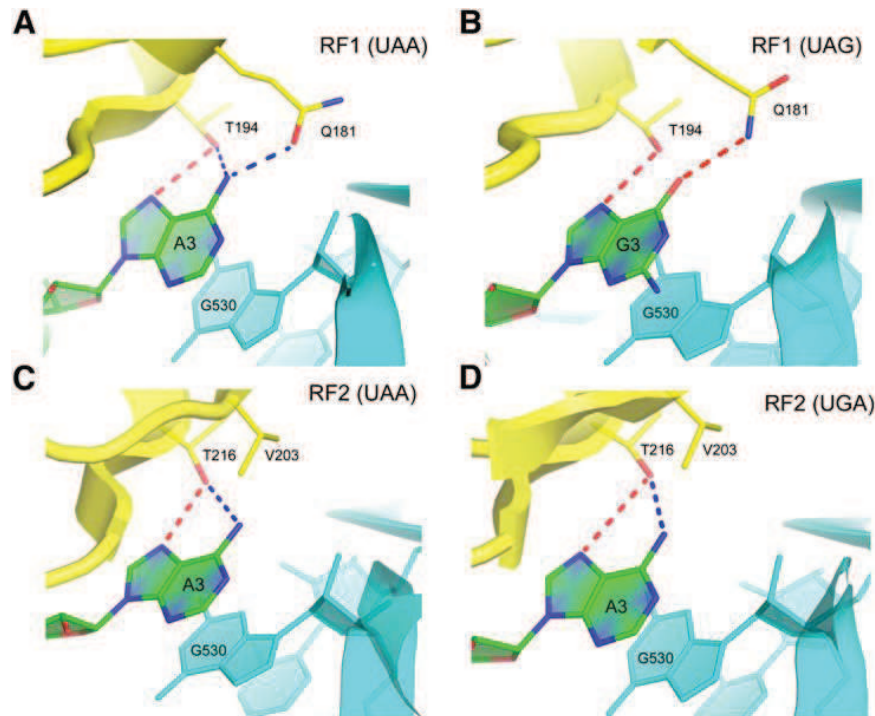
Figure 14: Interactions of the third nucleotide of the stop codon with the release factor. Adapted from (Korostelev, 2011).

Third base recognition in stop codon (A3/G3)

The third base recognition is the most remarkable, since the codon conformation is altered on binding RF such that the third nucleotide of the stop codon is unstacked from the first two and is recognised separately, in the G530 pocket. A conserved Thr residue at the C terminal end of the recognition loop examines the Hoogsteen edge of the third nucleotide and forms H-bonds with A3/G3 (Fig. 14). Gln181 present at the N-terminal end of the recognition loop is required by RF1 to allow A3 and G3 recognition. It is unusual, as it can accept as well as donate H-bond depending upon the nucleotide present. This Gln residue is replaced by a Val in RF2, which would make it incapable of forming H bond with G3 and thus make it specific for A3 (Korostelev et al., 2008b).

1.10 Peptidyl tRNA hydrolysis

The PTC protects the peptidyl tRNA in a compact pocket; formed by A2451, C2452, U2506, U2585 of 23S rRNA and A76 of tRNA, against premature hydrolysis (Schmeing et al., 2005a). However, on recognising a stop codon it must allow peptide release. The universally conserved GGQ motif in domain 3 of class I RF contacts the nucleotides of 23S rRNA and P-site tRNA. The PTC conformation with RF, as determined from crystal structures, is similar to that observed for ribosome structures with tRNAs bound in the A- and P-sites (Schmeing et al., 2005a). U2506 and U2585 are retracted from the A-site binding pocket in the presence of aminoacyl tRNA or the RF (Voorhees et al., 2009). Initially, the side-chain of Gln residue of eRF1 was proposed to be involved directly in catalysis (Song et al., 2000; Trobro and Aqvist, 2009). However, in crystal structures this side-chain was observed to point away from the scissile ester bond (Korostelev et al., 2008a; Laurberg et al., 2008). Also, biochemical and mutational studies ruled out the possibility of the Gln side-chain being involved in catalysis. Rather, the Gln side-chain appears to contribute to the affinity of RF binding to the ribosome and probably increase specificity of the reaction by excluding nucleophiles other than water to allow hydrolysis (Shaw and Green, 2007).

Moreover, within the GGQ motif, instead of the Gln residue, the presence of Gly was found to be more critical. The substitution of Gly had a severe effect on peptide hydrolysis (10^4 fold slower rate of peptide release) (Mora et al., 2003; Zavialov et al., 2002). Similarly, substituting the Gln by Pro led to complete loss of peptide hydrolysis (Korostelev et al., 2008a). It is because the main-chain amide group of Gln forms a hydrogen bond with the 3'-OH of A76 (Laurberg et al., 2008; Santos et al., 2013). This is the reason why any mutation that alters the GGQ motif backbone conformation drastically affects catalytic activity. Such an involvement of the backbone amide group in catalysis is not exactly unique to the ribosome. It has been observed to stabilise transition state in proteases, esterases and GTPases (Maegley et al., 1996; Wilmouth et al., 2001) and might represent a conserved mechanism.

Hence, most likely, the GGQ motif plays a catalytic role by stabilizing the leaving group or the transition state intermediate as well as the hydrolysis product (Jin et al., 2010). With these studies, a plausible mechanism of coordinating water molecule for peptide hydrolysis has been postulated (Korostelev, 2011) (Fig. 15).
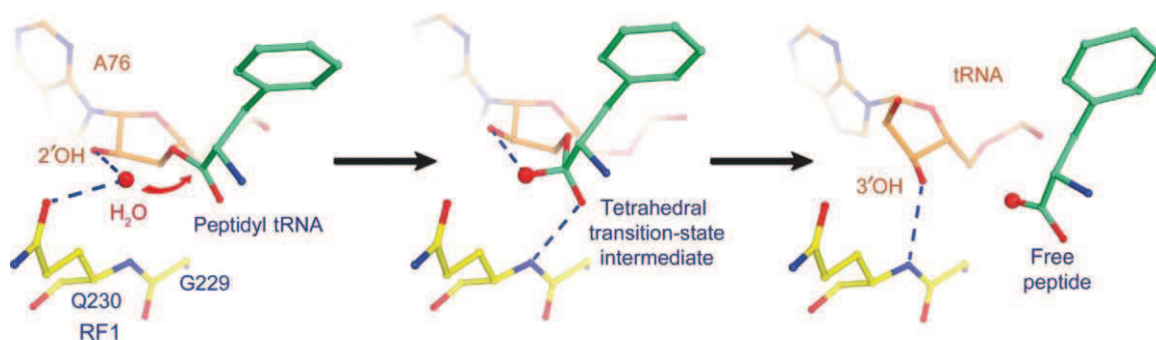
Figure 15: Postulated mechanism for peptidyl tRNA ester bond hydrolysis. A water molecule is positioned for a nucleophilic attack on the scissile ester bond. The middle panel depicts the tetrahedral intermediate state stabilization by the amide backbone of the essential 'GGQ' motif. In the last panel, the 3'-hydroxyl leaving group forms hydrogen bonds with the GGQ backbone post hydrolysis. Adapted from (Korostelev et al., 2010).

1.11 Coordination between the decoding centre and PTC for hydrolysis

The class I RF not only recognise the stop codon, instead they must somehow communicate with the PTC to allow peptide release. Considering that even in the absence of proofreading, the error rate of peptide release is as low as in the elongation step (10^{-3}-10^{-6}), there must exist a tight regulation between the two active sites on ribosome (Freistroffer et al., 2000; Wohlgemuth et al., 2010).

Structural studies have shown that RF1 adopts two different conformations, a "compact" and an "open" one and the loop connecting domains 3 and 4 (switch loop) might act as a switch between these two conformations (Vestergaard et al., 2005). The open conformation, most likely, represents the catalytically active conformation where the SPF and GGQ motifs are placed apt distance (73 Å) apart to allow decoding and peptide hydrolysis, thereby bridging the two active sites on the ribosome (Ma and Nussinov, 2004). Moreover, this switch loop is positioned in a pocket in the decoding centre formed by rpS12, A1492, A1493 (h44) and A1913 (h69) which were shown to participate in sense codon discrimination by aminoacyl tRNAs (Laurberg et al., 2008). In cognate tRNA complexes A1492 and A1493 of h44 are flipped out. Instead, A1492 and A1913 flip out on RF binding, while A1493 remains stacked, making interactions with A1913. This pocket formation on the ribosome is coupled with stop codon recognition in the decoding centre, which in turn places the switch loop in vicinity of h69 and positions the GGQ motif in the PTC to allow

peptide hydrolysis (Korostelev et al., 2010; Laurberg et al., 2008; Weixlbaumer et al., 2008). Thus, these residues are not directly involved in stop codon recognition and instead appear to coordinate peptide hydrolysis with stop codon recognition. In agreement with these results, disruption of interactions between helix 69 and the switch loop results in 10^3 fold decrease in the rate of peptide release (Korostelev et al., 2010).

Also kinetic studies have shown that the rate with which the RF1 associates with the ribosome (k_{on}) does not depend on the presence of a stop codon. Conversely, the rate of dissociation of RF1 (k_{off}) in the presence of stop codon is drastically lowered (Hetrick et al., 2009). Thus, stop codon recognition is the limiting factor rather than RF1 association for termination.

Based on these observations it has been proposed that initially, the class I RF binds to the ribosome in an inactive compact conformation. On recognising a stop codon in the decoding site by the reading head, a conformational change (in the switch loop) would allow docking of GGQ motif into the PTC and catalysis (Korostelev, 2011). This would lead to peptide release, thereby concluding the role of RF1/RF2.

1.12 RF3 structure

Gao and co-workers determined the crystal structure of RF3 which exhibits a three-domain architecture of RF3 (Gao et al., 2007) (Fig. 16A). Domain I is subdivided into the GTPase domain and an "EF-G like" G'-subdomain. The classical G-domain is composed of a six-stranded β-sheet (five parallel and an antiparallel β-strand) lined by six α-helices and a 3_{10}-helix, and is present in all the four major GTPases involved in translation; IF2, EF-G, EF-Tu and RF3 in prokaryotes and eIF2, eEF2, eEF1a and eRF3 in eukaryotes.

Domain II forms a β-barrel structure, which also exists as a fold in EF-Tu, EF-G and eRF3 polypeptides. A short peptide linker connects domain II to domain III. Domain III comprises another β-barrel along with two α-helices on either side.
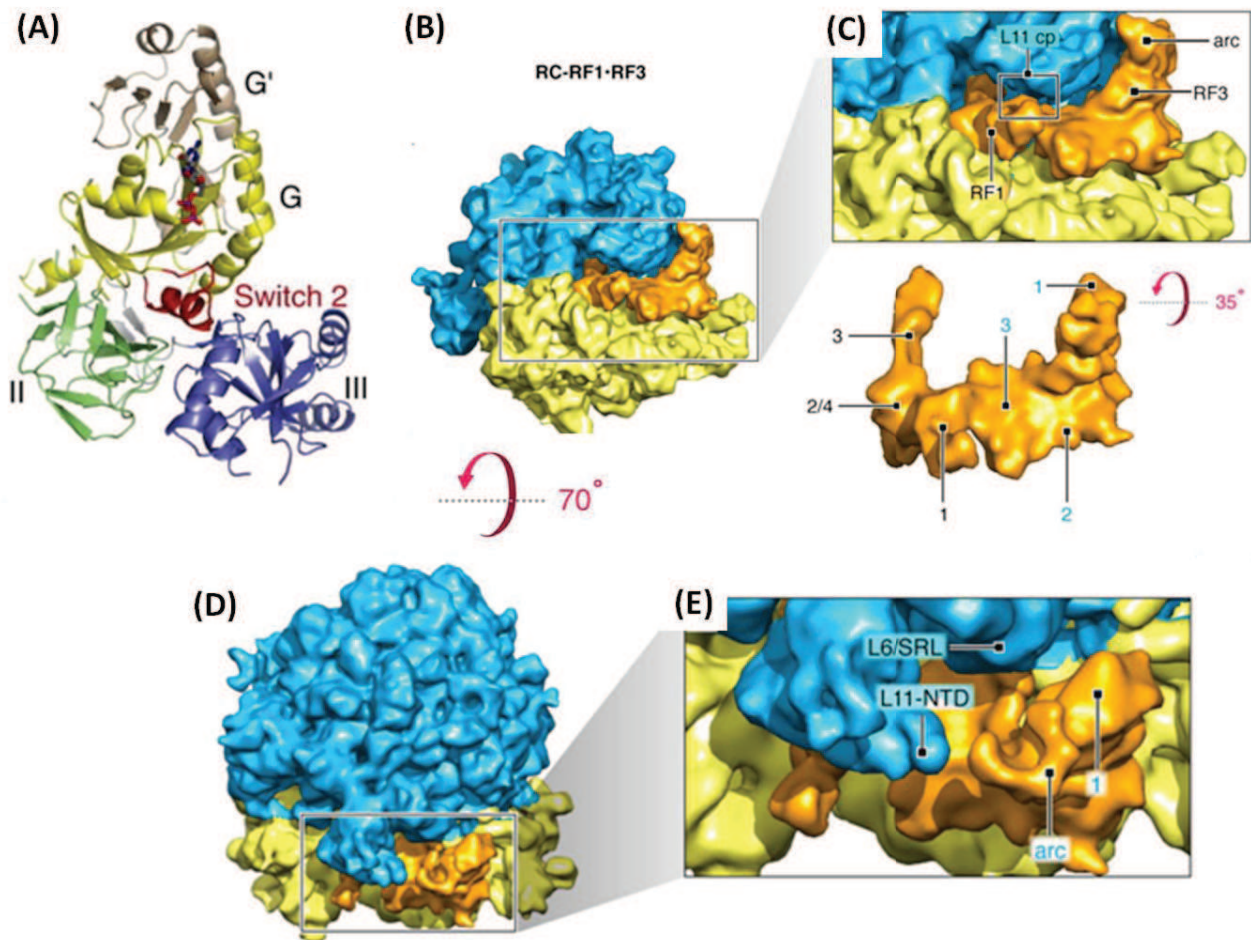
Figure 16: Class II RF. (A) crystal structure of RF3 with G, G', domains, domain II and III marked in light brown, yellow, green and blue respectively. Switch loop is marked in red. (B) Cryo-EM structure of RF1-apo-RF3 bound to 70S. 50S is in blue and 30S in yellow. (C) Zoomed-in view showing RF1 contacts with the ribosome at the L7/L12 stalk. Below is the density corresponding to RF1 and RF3, with domains marked in black and blue, respectively. (D) Cryo-EM structure from (C) rotated by 70°, showing apo-RF3 and its lack of interaction with L7/L12 stalk (E). The arc is the L7/L12 stalk. Adapted from (Gao et al., 2007) and (Pallesen et al., 2013) with some changes.

1.12.1 The role of class II release factors

Post peptide release, the 70S ribosome still has tRNA, mRNA and RF1/RF2 bound to it. It must release these components and separate out the two subunits to start another round of translation. This is where the class II RF come in picture (confirmed only for prokaryotes). RF3 binds in the GDP state to the RF1/RF2 bound ribosome. The exchange of GDP for GTP, releases the class I RF and RF3 dissociates on GTP-hydrolysis.

Recently, two crystal structures of RF3 (bound to GTP analogue) with the ribosome (*E. coli* and *T. thermophilus*) (Jin et al., 2011; Zhou et al., 2012) confirmed that it binds at the GAC in the small subunit, similar to EF-G binding site. The small subunit undergoes an

intersubunit rotation of about 7° and the head is also rotated counter-clockwise by 14° (ratcheting movement). Also, the 50S subunit is not free of conformational changes. The L1 stalk in the 50S subunit flexes towards the central protuberance (Gao et al., 2007). These changes alter the inter-subunit bridges. The RNA-rich bridges near the axis of rotation at bridge B3 are preserved, while the peripheral bridges like B1a (between protein S13 and h28 of 23S rRNA) and B1b (between protein S13 and L5) are disrupted and replaced by R1b between S13 and S19 from 30S and L5 from 50S.

These structures also shed light on the position of RF3 on the ribosome. It binds such that domain I contacts the 50S at the sarcin-ricin loop of 23S rRNA and protein L6, as observed for G domains of other GTPases. 30S contacts are made by domains II and III, at helices h5 and h15 of 16S rRNA and protein S12. RF3 itself also changes conformation on binding to the ribosome. Both domains II and III are rotated as compared to the free RF3 crystal structure. Additionally, the switch loop I in GTP binding domain is ordered as compared to the crystal structure and forms an enclosure around the GTP analogue, probably permitting GTP hydrolysis.
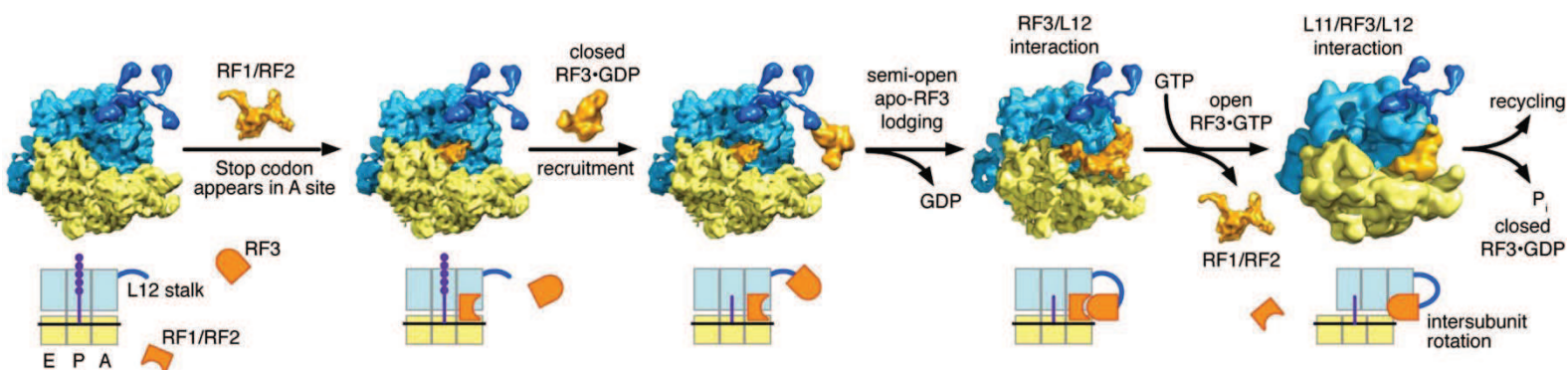
Figure 17: Proposed model for translation termination. A class-1 RF recognises its cognate mRNA stop codon in the ribosome, binds in the A-site and mediates release of the nascent protein attached to the P-site tRNA. Post peptide release, RF3•GDP is recruited to the ribosome in its closed form probably mediated by L12-CTD and does not form a stable complex with the ribosome. As RF3 accommodates onto the ribosome, GDP is released and apo-RF3 assumes its semi-open conformation contacting L12-CTD, the class-1 RF and 30S protein S12. On binding GTP, RF3•GTP assumes its open conformation and the ribosome changes from the unrotated to the rotated conformation. This allows the class-1 RF to leave and GTP hydrolysis to occur in the complex probably due to contact with L6/SRL and L12-CTD. RF3•GDP dissociates from the ribosomal complex in its closed conformation and the ribosome is ready for subunit recycling. Adapted from (Pallesen et al., 2013).

However, until last year, there had been no structures of both the release factors bound simultaneously to the ribosome. This missing link was partially sealed by the cryo-EM structure of RF1 and apo-RF3 bound to 70S (Pallesen et al., 2013). It showed that apo-RF3 in presence of RF1 interacts differently than if only RF3 is present. Domain 1 of apo-RF3 does not contact the SRL, instead it is just present in its vicinity on the ribosome (Fig. 16). The apo-RF3 is anchored onto the ribosome solely through domain III contacts. Moreover, apo-RF3 interacts with the L7/L12 stalk which was not the case for RF3 bound with nucleotide. Thus, in order to proceed from apo-RF3 to RF3 with GTP, the ribosomal L7/L12 stalk would undergo an upward rotation as depicted in the figure 16. Based on these observations, (Pallesen et al., 2013) have proposed a mechanism for translation termination (Fig. 17).

1.13 Analogy between termination and transpeptidation

The two processes of transpeptidation and peptidyl-tRNA hydrolysis are similar and dissimilar at the same time. The components that catalyse both the processes are similar in structure, RF3 and EF-G (Fig. 18). Also, both of these entities depend upon codon recognition in the A site to stimulate conformational changes and nucleophilic reaction in the PTC. The nucleophile in transpeptidation is the α-amino group of the aminoacyl tRNA in A site whereas for the termination reaction it is an incoming water molecule. But, both the nucleophiles attack the carbonyl carbon atom of the ester bond.
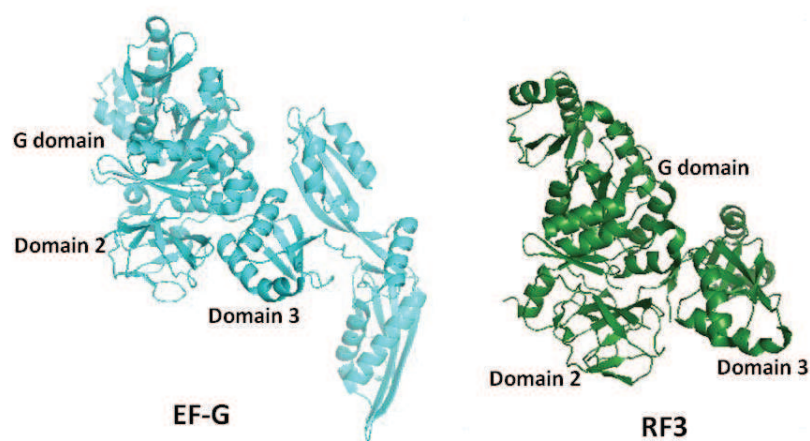
Figure 18: Structural similarity between EF-G (blue, PDB code 2DY1) and RF3 (green, PDB code 3VQT). The homologous G domain and domains 2 and 3 are marked.

In addition, a proton shuttle mechanism, like as employed during elongation, has been suggested for termination (Korostelev et al., 2010; Rodnina et al., 2006). The conformational changes in the ribosome involving ratchet like rotation of the small subunit, swivelling of the L1 stalk and the movement of the deacylated tRNA from P-site to the hybrid P/E state are observed on binding of the RF3 as well as EF-G.
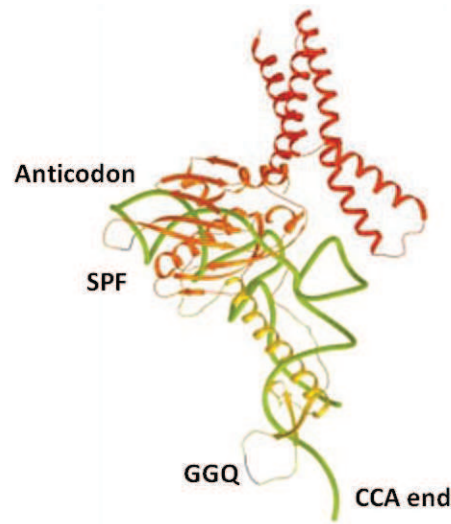
Figure 19: Functional similarity between RF2 and tRNA is not exactly corroborated by structural similarity. tRNA (green) superposed on RF2. Domain 1 of RF2 is shown in red, domain 2/4 in orange, and domain 3 in yellow. GGQ and anticodon ends are marked. Adapted from (Klaholz et al., 2003).

However, that is where the similarity ends. The class I RFs and aminoacyl tRNAs that recognise the respective codons, have different structures (Fig 19). Furthermore, unlike the peptidyl transfer reaction, peptidyl tRNA hydrolysis requires direct participation of a class I release factor. Secondly, the codon recognition mechanisms are different for binding tRNA or release factor. Also, the proof reading mechanism which is essential for elongation has not been found in bacterial stop codon recognition (Freistroffer et al., 2000). Finally, the outcome of the two processes is different as the aminoacyl tRNA binding leads to a transpeptidation reaction, and instead the stop codon recognition allows hydrolysis of the peptidyl tRNA.

1.14 Eukaryotic translation termination

The termination in eukaryotes attains a higher level of complexity and regulation as compared to prokaryotes. With higher complexity come the inherent challenges to understand the mechanisms. The factors involved are well studied but the interaction with the ribosome and the mechanism of stop codon recognition are not understood at the molecular level.

1.15 Structure of eRF1

RF1 and eRF1 in spite of having the same function, share no sequence or structural homology. The crystal structure of human eRF1 showed that it consists of three domains (Song et al., 2000), resembling a "Y" shape (Fig. 20 B). The N-terminal domain (NTD) is composed of a four stranded β-sheet, enclosed on both sides by two α-helices. The TASNIKS and YxCxxxF motifs, involved in stop codon recognition are present in this domain (Bertram et al., 2000; Frolova et al., 2002). The TASNIKS motif is placed between helices α2 and α3. The helix α1 creates an interface with domain 3.

The M domain forms the stem of the "Y" shape. The universal GGQ motif, essential for peptidyl-tRNA hydrolysis is present in a loop at the tip of the stem. A long continuous α-helix connects domains 2 and 3. The C-terminal domain is similar to an α/β sandwich fold. Although the majority of the residues were not characterised in the crystal structure, NMR studies have shown that a minidomain is formed by an α-helix and three β-strands in the C domain (Mantsyzov et al., 2010). This C-domain interacts with the C-terminal domain of eRF3 (Cheng et al., 2009). Thus, none of these domains bear any resemblance to the 4 domains of RF2. Also, unlike the compact form of RF1 crystallisation, eRF1 crystallises in an extended form.

eRF1 must undergo a conformational change on binding eRF3 so that the two catalytically active regions can be apt distance apart to allow stop codon recognition, and peptide hydrolysis (Cheng et al., 2009). In this bent conformation, it resembles a tRNA molecule. Functionally, eRF1 is also involved in the next translational phase of recycling. After peptide release, it remains bound on the ribosome and coordinates with ABCE1 (ATP-

binding cassette sub-family E member 1) to promote 80S splitting into free 60S, 40S, mRNA and tRNA (Pisarev et al., 2010).
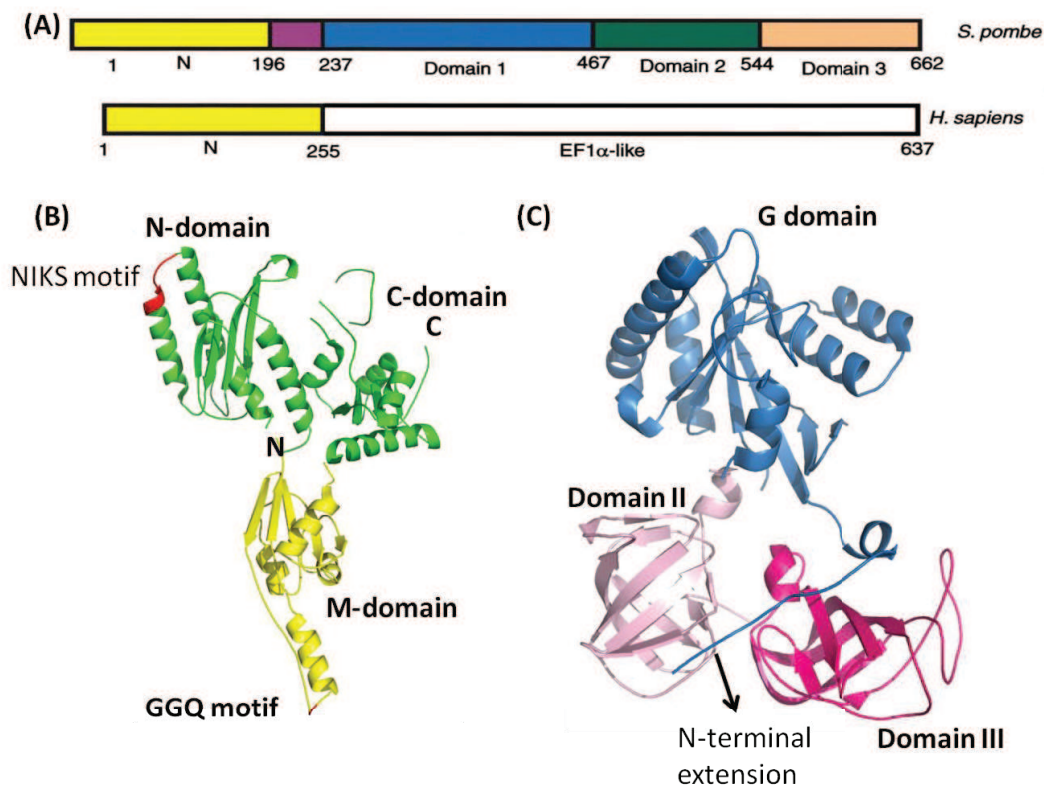
Figure 20: (A) Schematic representation of eRF3 for *S. pombe* and humans. (B) eRF1 crystal structure with catalytically active NIKS and GGQ motifs marked in red N-domain (green) and M-domain (yellow). PDB code 1DT9. (C) eRF3 crystal structure without N-terminal domain with domain 1, 2 and 3 in blue, light and dark pink respectively. PDB code 1R5B.

1.16 Structure of eRF3

The structure of *S. pombe* eRF3 lacking the N-terminus (1-196), sharing 52% sequence similarity with human eRF3, is the most complete structure of eRF3 available till now (Fig. 20). The polypeptide chain of eRF3 from *S. pombe* is broadly divided into a N-terminal non homologous region and a conserved C-terminal region resembling the fold of eEF1α. The latter is further divided into a three-domain architecture. Domain 1 is the conserved GTPase domain with a six stranded β-sheet (5 parallel and 1 antiparallel β-strand) lined by six α-helices and a 3_{10}-helix. Domains 2 and 3 each assume a β-barrel structure as observed for eEF1α, EF-Tu and human eRF3 (Cheng et al., 2009; Kjeldgaard and Nyborg, 1992; Song et al., 1999) (Fig. 21). The last two domains are indispensible for interactions between eRF1 and eRF3 while the GTPase domain is not necessary for such interactions.

The structure of the N-terminal domain of eRF3 has not yet been determined. It contains a long poly G stretch at the beginning of the N-domain. But functional studies have shown that it is not involved in GTP-hydrolysis or eRF1-eRF3 interactions (Kushnirov et al., 1988).
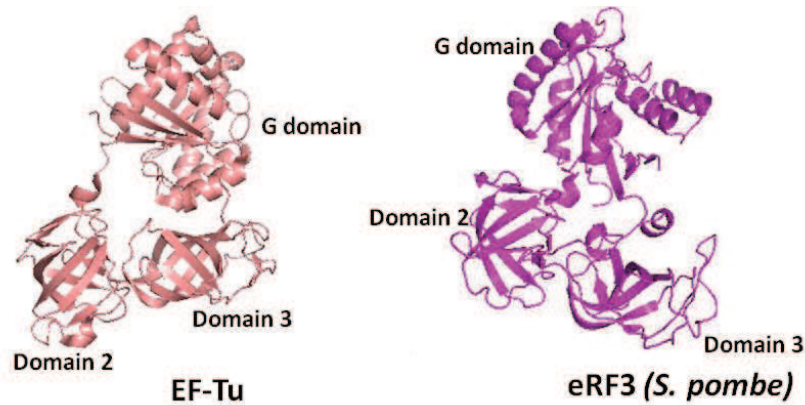
Figure 21: Structural similarity between EF-Tu (salmon red, PDB code 4J0Q) and eRF3 (magenta, PDB code 1R5B). The homologous G domain and domains 2 and 3 are marked.

Also, the role of eRF3 is completely different as compared to RF3. RF3 mediates recycling of RF1/RF2 after peptide release (post termination). On the contrary, translation termination *per se* is highly dependent on the presence of both factors in eukaryotes. GTP-hydrolysis by eRF3 allows coordinating eRF1's codon recognition and peptide release activities, as explained below.

Functionally, eRF3 is involved in several processes apart from translation termination. It has been recently shown to undergo self cleavage at Ala73 to yield a processed isoform (p-eRF3) which localises in the nucleus as well as cytoplasm. p-eRF3 promotes apoptosis by interacting with inhibitors of apoptosis proteins and releasing the caspases and might be involved in regulating cell death (Hashimoto et al., 2014). In addition, eRF3 has been hypothesised to ensure translation continuation, linking termination with initiation. It binds to PABPC1 (cytosolic poly-A binding proteins) via two PAM2 motifs present in the N-domain (Kononenko et al., 2010). PABPC1 is bound to the 3' poly-A tail of the mRNA. These PAM2 motifs (PABPC1-interacting motif 2), PAM2-N and PAM2-C, residues 67–78 and 76–87, respectively, interact with cytosolic polyA binding domain of PABPC1 (Osawa et al., 2012). The PABPC1, in turn is known to interact with eIF4G which might be involved in forming a closed loop wherein the 3'UTR is looped out, and the mRNA 5'-cap is connected with ribosomes that are involved in termination.

Furthermore, eRF3 mediates mRNA decay by participating in mRNA poly-A tail degradation. The enzymes Pan2–Pan3, Caf1–Ccr4 (Uchida et al., 2004); required for deadenylation possess the PAM2 motifs, and bind to PABPC1 only on release of eRF3 from PABPC1 post translation (Funakoshi et al., 2007). The interaction of eRF3 with PABPC1 is thought to be a mechanism to protect the mRNA from being deadenylated. Thus, eRF3 also actively ensures mRNA turnover in the cell.

```
eRF3a           --MDPGSGGGGGGGGGGGSSSGSSSSDSAPDCWDQADMEAPG----------------- 40
eRF3b           --MD---------------SGSSSSSDSAPDCWDQVDMESPG----------------- 24
eRF3-pombe      -MASNQPNNGEQDEQLAKQTSKLSMSAKAPTFTPKAAPFIPS-----------FQRPGF 47
eRF3-cerevisiae MSDSNQGNNQQNYQQYSQNGNQQQGNNRYQGYQAYNAQAQPAGGYYQNYQGYSGYQQGGY 60
                     .           .  . .          *.

----------- ⚡ ------------------------------- ⚡ -------------------------------------
eRF3a           MEEEEEIPKPKSVVAPPGAPKKEHVNVVFIGHVDAGKSTIGGQIMYLTGMVDKRTLEKYE 247
eRF3b           MEEKEEIRKSKSVIVPSGAPKKEHVNVVFIGHVDAGKSTIGGQIMFLTGMVDKRTLEKYE 239
eRF3-pombe      TDLQN---EVDQELLKDMYG-KEHVNIVFIGHVDAGKSTLGGNILFLTGMVDKRTMEKIE 274
eRF3-cerevisiae ALIKEQEEEVDDEVVNDMFGGKDHVSLIFMGHVDAGKSTMGGNLLYLTGSVDKRTIEKYE 296
                ::    : .. :        *:**.::*:*********:**:::*** ***** :** *

eRF3a           REAKEKNRETWYLSWALDTNQEERDKGKTVEVGRAYFETEKKHFTILDAPGHKSFVPNMI 307
eRF3b           REAKEKNRETWYLSWALDTNQEERDKGKTVEVGRAYFETERKHFTILDAPGHKSFVPNMI 299
eRF3-pombe      REAKEAGKESWYLSWALDSTSEEREKGKTVEVGRAYFETEHRRFSLLDAPGHKGYVTNMI 334
eRF3-cerevisiae REAKDAGRQGWYLSWVMDTNKEERNDGKTIEVGKAYFETEKRRYTILDAPGHKMYVSEMI 356
                ****:  .:: *****.:*:..***:.***:***:*****::::::******* :*.:**

eRF3a           GGASQADLAVLVISARKGEFETGFEKGGQTREHAMLAKTAGVKHLIVLINKMDDPTVNWS 367
eRF3b           GGASQADLAVLVISARKGEFETGFEKGGQTREHAMLAKTAGVKHLIVLINKMDDPTVNWS 359
eRF3-pombe      NGASQADIGVLVISARRGEFEAGFERGGQTREHAVLARTQGINHLVVVINKMDEPSVQWS 394
eRF3-cerevisiae GGASQADVGVLVISARKGEYETGFERGGQTREHALLAKTQGVNKMVVVVNKMDDPTVNWS 416
                .*******:.*******:**.*:***.********:**:* *:::::*::****.*:*:**

eRF3a           NERYEECKEKLVPFLK-KVGFNPKKDIHFMPCSGLTGANLKEQSD--FCPWYIGLPFIPY 424
eRF3b           IERYEECKEKLVPFLK-KVGFSPKKDIHFMPCSGLTGANIKEQSD--FCPWYTGLPFIPY 416
eRF3-pombe      EERYKECVDKLSMFLRRVAGYNSKTDVKYMPVSAYTGQNVKDRVDSSVCPWYQGPSLLEY 454
eRF3-cerevisiae KERYDQCVSNVSNFLR-AIGYNIKTDVVFMPVSGYSGANLKDHVDPKECPWYTGPTLLEY 475
                ***.:*  .:: **:    *:. *.*: :** *. :* *:*:::  *   **** * .:: *

------------- ⚡ ------------------------------- ⚡ -------------------------------------
eRF3a           TFKDFPQMGRFTLRDEGKTIAIGKVLKLVPEKD 636
eRF3b           TFKDFPQMGRFTLRDEGKTIAIGKVLKLVPEKD 628
eRF3-pombe      RFEDYQYMGRFTLRDQGTTVAVGKVVKILD--- 662
eRF3-cerevisiae TYQDYPQLGRFTLRDQGTTIAIGKIVKIAE--- 685
                ::*:   :*******:*.*:*:**::*:
```

N-terminus

G-domain

Domain III

Figure 22: Above is the ClustalW (Goujon et al., 2010) sequence alignment of eRF3 from humans and yeast. The deleted sequence is separated by (-⚡-- ⚡ -). eRF3a, eRF3b sequences are from humans while third sequence of eRF3 is from S. pombe and the last sequence is from S. cerevisiae. In red are the highly conserved sequences and in blue is the poly-G sequence at the N terminus of eRF3a which causes polymerase slippage during cloning.

Table 3: ClustalW sequence alignment score for the above mentioned eRF3 sequences.

Seq A	Name	Seq B	Name	Similarity score
1	eRF3a	2	eRF3b	85.9873
1	eRF3a	3	eRF3-pombe	37.1069
1	eRF3a	4	erf3-cerevisiae	41.6667

The eRF3 identified in humans, mouse and rat genomes has two variants; eRF3a and eRF3b, encoded by GSPT1 and GSPT2 genes respectively, sharing long non homologous stretches at their N termini. Fig. 22 and Table 3, indicates their homology and sequence alignment using ClustalW (Goujon et al., 2010). The eRF3a gene possesses several intronic sequences, and its mRNA is present ubiquitously in all mouse tissues with varying expression levels during the cell cycle. In contrast, the eRF3b gene has no intronic sequence and is poorly expressed in most tissues except brain. Both the variants can interact with eRF1 and stimulate its activity *in vitro* efficiently (Jakobsen et al., 2001; Zhouravleva et al., 1995). The full-length human eRF3 structure is not yet determined. However, there have been structures of human eRF1 with truncated eRF3 (lacking N terminus) in ternary complexes with and without ribosomes, as explained below.

1.17 Human eRF1-eRF3 complex and inter-dependability

Unlike in prokaryotes, the two release factors in eukaryotes form a stable complex in solution (Ebihara and Nakamura, 1999; Ito et al., 1998) and function in an interdependent manner. In mammalian cells, eRF1 alone is enough to allow efficient termination which might lead us to conclude that eRF3 is dispensable (Frolova et al., 1994), but in yeast, eRF3 is a prerequisite for viability (Stansfield et al., 1995). However, the very existence of eRF1-eRF3 protein complex in eukaryotic cells is intriguing, but its purpose remains obscure.

The termination of protein synthesis on a ribosome requires the two proteins to work in an interdependent manner. eRF3's GTP binding activity has been determined to be highly dependent on eRF1 (Hauryliuk et al., 2006; Pisareva et al., 2006). eRF3, in turn, stimulates the peptide release activity of eRF1 (Alkalaeva et al., 2006). Also, the GTPase activity of eRF3 requires eRF1 presence on the ribosome, whereas the prokaryotic RF3 does not require RF1/RF2 for GTP hydrolysis (Frolova et al., 1996).

The crystal structure of human eRF1-eRF3 protein complex contains full length eRF1 and truncated eRF3 lacking the N-terminus and G-domain. The eRF1 C-terminal domain interacts with domain 3 of eRF3 mainly through van der Waals and hydrophobic contacts. These hydrophobic patches on eRF3 are highly conserved in yeast and humans, like 'GRFTLRD' motif. But SAXS analysis has shown that M-domain of eRF1 contacts the GTPase

domain of eRF3 (Cheng et al., 2009) (Fig. 23). The presence of this M-domain is absolutely crucial for GTP-binding and hydrolysis by eRF3 (Kononenko et al., 2008). This is corroborated, to some extent, by cryo-EM studies, as explained below. But the eRF1-eRF3 complex with ribosome still needs to be studied in detail to shed light on termination mechanism.
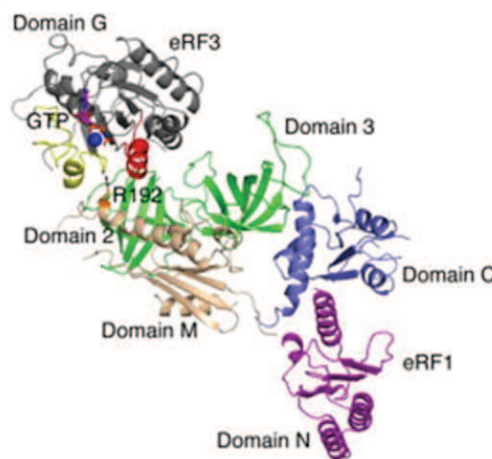
Figure 23: SAXS model for eRF1-eRF3-GTP complex. Adapted from (Cheng et al., 2009).

1.18 Interactions with mammalian ribosome

Crystal structures for individual factors and complexes have been available but the limiting factor for crystal structures of eukaryotic termination complexes with ribosomes is the crystallization process itself for higher eukaryotic ribosomes. So far, the insights into the eukaryotic translation termination mechanism have been from biochemistry and cryo-EM studies. In the last two years structures of the eRF1-eRF3 associated mammalian termination complex (des Georges et al., 2014; Taylor et al., 2012) and Dom34-Hbs1, Pelota-Hbs1 associated recycling complexes (Becker et al., 2012) have brought to light the communication between factors and ribosomes (Fig. 25).

eRF3 was observed to bind at the universal GAC on the 60S with its G domain contacting the sarcin-ricin loop (H95) of 28S rRNA. It is in close proximity of the P stalk, rpL40e and rpL9. The domain 2 β-barrel structure contacts h5 and h15 of 18S rRNA in the 40S (Fig. 24). The G domain and domain 2 sandwich the M-domain of eRF1, placing the GGQ motif, present at the tip of the M-domain, in the intersubunit space of the 80S. Also, in the immediate vicinity of the M-domain is the switch-I region of eRF3, which might be stabilised

on interaction with eRF1 enabling GTP-binding to eRF3. Fig. 24 shows a structural comparison of *S. pombe* eRF3 and cryo-EM ribosome bound form of eRF3.
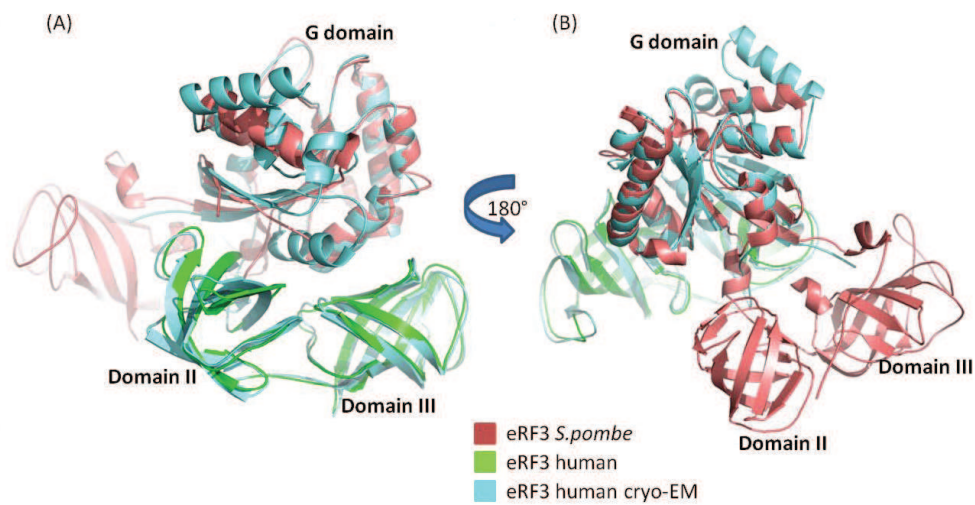
Figure 24: Structural comparison of *S. pombe* eRF3 crystal structure (salmon red), human eRF3 crystal structure (green) and human eRF3 cryo-EM structure when bound to ribosome. (A) Shows the overlap in structures of the G domain from both the species. (B) Domain II and III have different conformations in cryo-EM and crystal structures.

As predicted by biochemical studies, the N-domain of eRF1 contacts the helices h18, h30, h31, h34, h44 of 18S rRNA and proteins rpS30e and rpS31e in the 40S decoding centre (Bertram et al., 2000; Frolova et al., 2002). The TASNIKS motif is placed in immediate vicinity of the stop codon, in accord with UV cross-linking studies (Chavatte et al., 2002). The YxCxxxF motif is also close to the stop codon and within interacting distances of bases in positions 1 and 2 (Fig. 24). And the C-domain of eRF1 contacts the P-stalk base of the 60S subunit via RPL12, H43 and H44. The minidomain contacts the beak of the small subunit and might be required to limit the intersubunit and head rotation to permit efficient peptide release. At the same time, the C-domain of eRF1 communicates with the domain 3 β-barrel of eRF3 as observed in the crystal structure and SAXS studies of free eRF1-eRF3 proteins (Cheng et al., 2009).

The remarkable aspect of these studies is the non-rotated state of the ribosome and the open position of L1 stalk. The P-stalk base of 60S has an altered conformation and so does the rpS3 of 40S. New head body connections are formed on the 40S leading to constriction of the mRNA entrance channel. Ribosome conformational changes on binding of eRF1-eRF3-GTP ternary complex to ribosomes were predicted (Alkalaeva et al., 2006),

since they observed a two nucleotide forward shift of ribosome toe print. These structural studies confirm this.
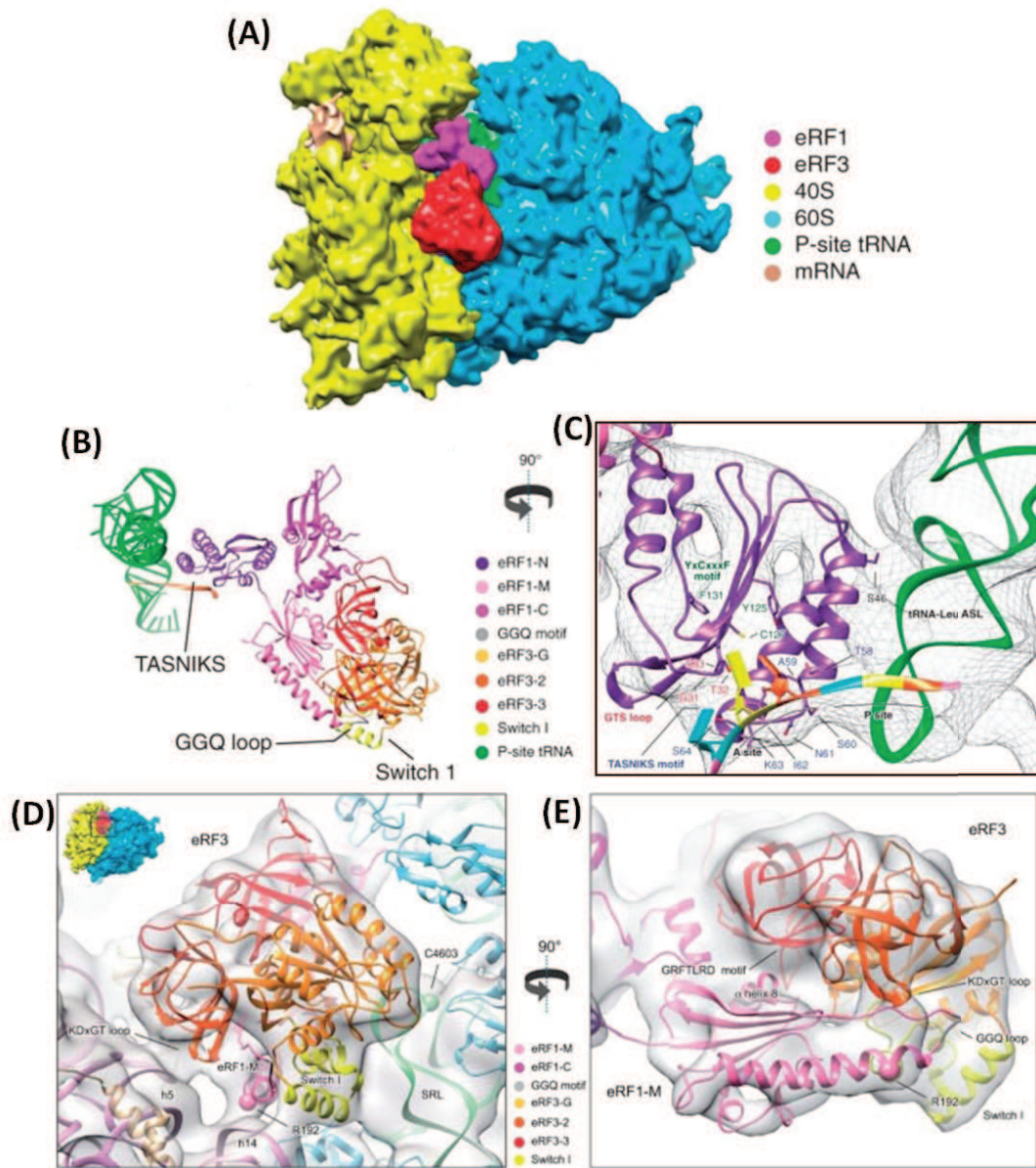
Figure 25: Pre-termination complex overview showing 40S, 60S, eRF1, eRF3, P-site tRNA and mRNA. (B) Atomic model fitted into the cryo-EM map showing eRF1, eRF3 and P-site tRNA. (C) Zoomed in view of eRF1 N-domain and mRNA showing the TASNIKS, YxCxxxF motifs and their positions relative to the modelled stop codon containing mRNA. First, second and third codon positions are indicated in orange, yellow and blue. (D)- (E) Close up of eRF3 depicting human 80S model fitted in the cryo-EM density map. 18S rRNA is depicted in purple, 40S proteins in beige and 60S proteins in light blue. Adapted from (des Georges et al., 2014).

Thus, the eRF1-eRF3-GMPPNP complex on the ribosome takes on a conformation similar to 70S-bound tRNA–EF-Tu and yeast 80S-bound Dom34–Hbs1 complexes. Considering that this complex has a non-hydrolysable analogue of GTP, GMPPNP bound, it

represents the pre-termination state of translation. The GGQ loop is placed far away from the PTC, indicating that GTP-hydrolysis is a prerequisite for accommodation to catalyse peptide hydrolysis. The GTP-hydrolysis would also change conformation of switch regions in G-domain so that eRF1 M-domain is accommodated to release peptide. This is achieved by a hinge movement of M domain as observed in these cryo-EM structures. However, eRF1 has been reported to have a lower k_d in the post termination complex (Pisarev et al., 2010), post GTP hydrolysis. Becker and his co-workers showed that eRF3 and ABCE1 share the same binding site on the ribosome (Becker et al., 2012). This might be a mechanism to link termination with recycling by keeping eRF1 bound until ABCE1 has been recruited to promote subunit dissociation.

These structures, however, have a limited resolution (8-10Å) which leaves a lot of questions unanswered. Like the long-standing dogma about factor association with ribosome as a ternary complex, still must be proven. They neither provide insight into the decoding mechanism nor the peptide release mechanism. Moreover, post termination, the mechanism leading to dissociation of the two factors to ensure recycling of the subunits for next round of translation needs to be explored structurally in more detail.

1.19 Recycling and ribosome rescue

Translation termination step is involved in regulating and communicating with other pathways as well, like mRNA quality surveillance and viral response. Aberrant mRNA translation can cause stalling of the translation machinery leading to unavailability of ribosomes and potentially dangerous by-products for the cell.

Three mRNA surveillance mechanisms have been reported in eukaryotes. Premature translation termination triggers nonsense-mediated decay (NMD) while slowed or halted translation due to rare codons or pseudoknot or RNA stem-loop structures activates the No-go decay (NGD). The last of these decay mechanisms, nonstop-decay (NSD), acts on arrested translation machinery with mRNA lacking a stop codon and have the poly-A tail translated into poly-Lys stretches.
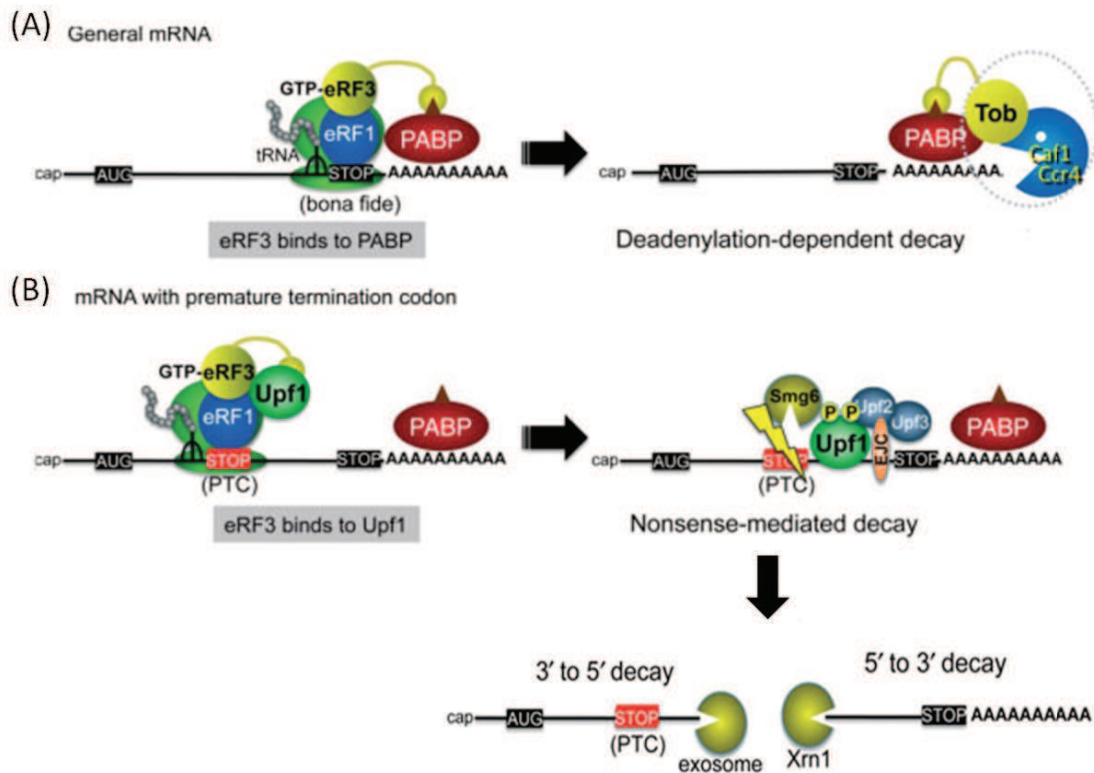
Figure 26: The role of termination factors in distinguishing an accurate stop codon from a premature stop codon. (A) eRF3 interacts with PABP at the canonical stop codon, triggering deadenylation dependent decay of the mRNA. (B) NMD pathway is activated on eRF3 interaction with Upf1 instead of PABP. Adapted from (Hoshino, 2012).

NMD depends on eRF1-eRF3 interactions with the specific factors like Upf1 (Fig. 26). Upf1, RNA-dependent ATPase/helicase, interacts with eRF3 and SMG1, a phosphatidylinositol kinase-related serine threonine protein kinase. eRF3 PAM2 motifs bind Upf1, so Upf1 must compete with PABPC1 to interact with eRF3. This depends on the length of mRNA. If mRNA is authentic, then eRF1-eRF3 complex will be able to interact with PABPC1 activating normal decay pathway. On the contrary, if mRNA possesses a nonsense codon, it will be unable to interact with PABPC1 and allow Upf1 binding to the translation machinery. Upf1, in turn will trigger its own phosphorylation by recruiting SMG1 and SMG6, thereby cleaving the nonsense codon encoding mRNA (Conti and Izaurralde, 2005) (Fig. 26).

In mammals, NGD and NSD pathways are more complicated and less well understood. ABCE1 (yeast homologue Rli1) is vital for dissociation of programmed and vacant/stalled ribosomes (Pisarev et al., 2010). eRF1 and eRF3 paralogues, Dom34 (yeast)/

Pelota (mammals) and Hbs1, respectively, coordinate with ABCE1 to promote dissociation of stalled elongation complexes in NGD and NSD (Doma and Parker, 2006; Tsuboi et al., 2012).

To understand these mechanisms in the context of termination, structures of the key players should be briefly described. Pelota (Dom34, yeast homologue, aPelota in archaea) possesses a 3-lobed structure, N-terminal, central and C-terminal domain, similar to eRF1. The motifs required for termination (NIKS, YxCxxxF and GGQ) are replaced by a patch of basic residues crucial for NGD. Hbs1 is comparable to EF-Tu like GTPases (including eRF3) comprising the G domain and β-barrels of domains II and III. The NTD is largely unstructured like for eRF3 but is variable in length. In archaea, aEF1α takes over the function of Hbs1.
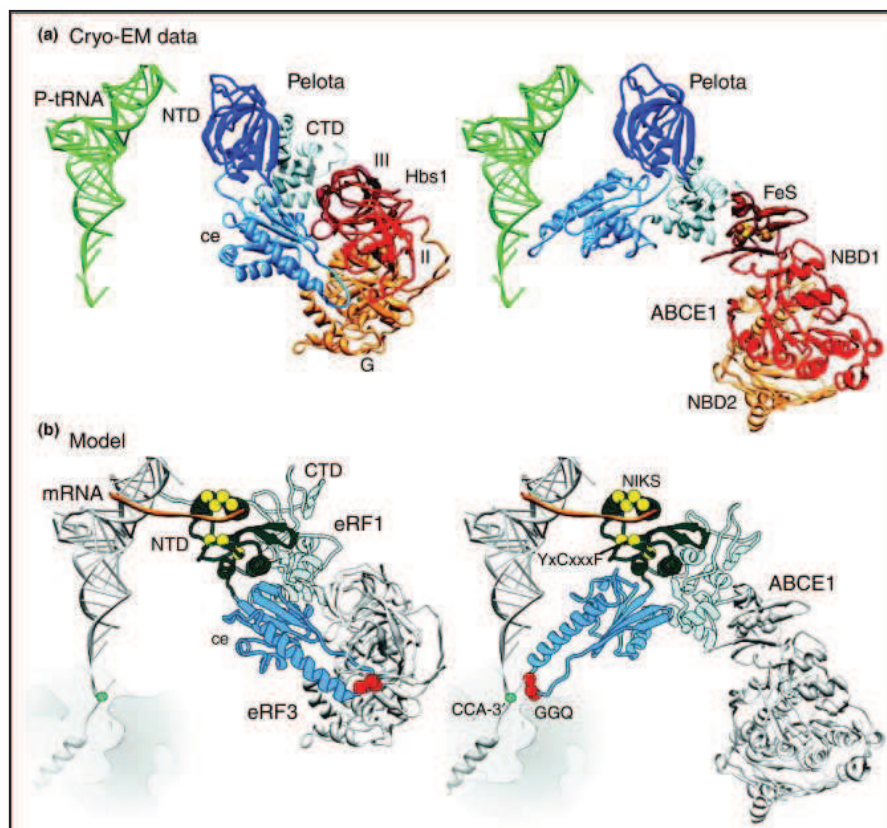
Figure 27: NGD pathway. (a) The central domain movement of Pelota as observed in cryo-EM structures of Pelota-Hbs1 and Pelota-ABCE1 bound to ribosome. (b) Model predicted for eRF1-eRF3 interaction and conformation change on ribosome, based on structures in (a). Adapted from (Franckenberg et al., 2012).

ABCE1 belongs to the ABC family of proteins characterised by two ABC-type nucleotide binding domains (NBD) arranged in a head-to-toe fashion. Each NBD has Walker A and Walker B motifs, A loop (with aromatic residues), Q, D, H loops (named so because of the conserved residues) and the ABC signature motif (LSGGQ) present in a α-helical domain.

ABC family of proteins have the ability to regulate access to NBD by conformational changes from an ADP bound open state to ATP bound closed state. Of all the family members, ABCE1 uniquely possesses an iron-sulphur domain (FeS) with two non-equivalent diamagnetic $[4Fe-4S]^{2+}$ clusters.

ABCE1 acts in coordination with an A-site factor on the ribosome like eRF1 or Pelota to split the complex into free 60S and tRNA, mRNA bound 40S. Kinetic studies have observed that ABCE1's ribosome recycling tendency is dependent on ATP-hydrolysis but is independent of peptide release. It stimulates eRF1 to release the peptide, but even in the absence of peptide release, ABCE1 could dissociate the two ribosomal subunits (Shoemaker and Green, 2011). In the cryo-EM structures of aPelota-aABCE1-ribosome and Pelota-Hbs1-ribosome (Becker et al., 2011), aABCE1 and Hbs1 bind at overlapping sites at the canonical GTPase centre. When Pelota is bound with Hbs1 on the ribosome, its central domain is tightly packed (Fig. 27). On swapping Hbs1 for ABCE1, Pelota undergoes a 140° rotation placing its central domain in close vicinity of the P-site tRNA acceptor stem and intersubunit bridge B2a formed by h44 and H69. Also, the Fe-S domain of ABCE1 interacts with Pelota rather than with the ribosome, and it might mediate this conformational change.
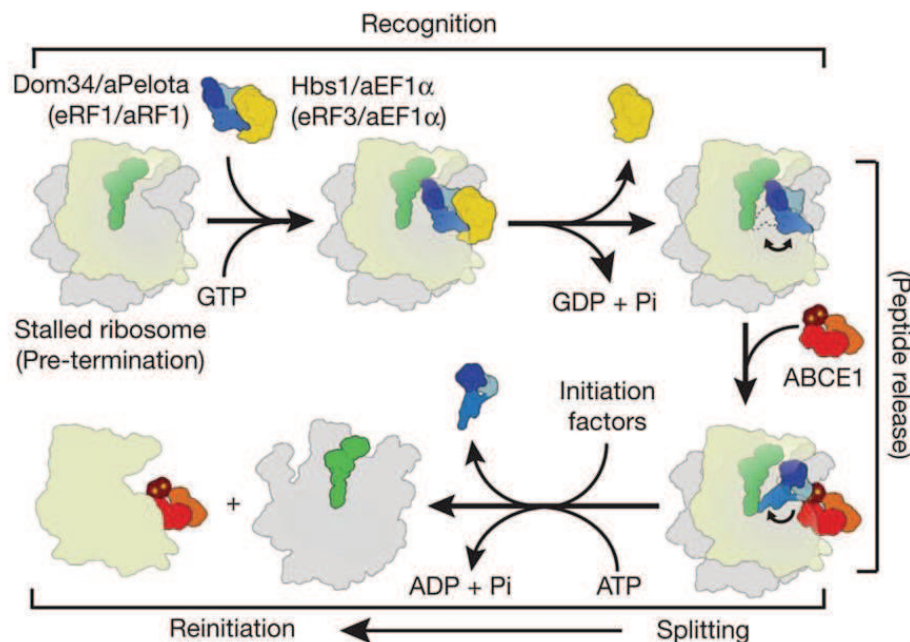
Figure 28: Proposed scheme for translation termination and ribosome recycling. Adapted from (Becker et al., 2012).

Considering the structural and functional similarity of Pelota and Hbs1 to eRF1 and eRF3, respectively, a model for termination mechanism has been hypothesised (Fig. 28). The eRF1-eRF3-GTP binds to the ribosome as a ternary complex, not individually as in prokaryotes. On binding to ribosome, the M-domain of eRF1 would be packed closely against eRF3, as depicted by SAXS studies. GTP hydrolysis would rearrange the M domain, placing the GGQ motif close to the 3'-CCA end of P-site tRNA to allow peptide release. This would be assisted by eRF3 liberation and ABCE1 recruitment. ABCE1 in turn would coordinate with eRF1 to ensure dissociation and turnover of components of the translation machinery (Fig. 28). This model, to some extent, has been corroborated by the cryo EM structure of pre termination complex, as discussed above.

1.20 Viral protein association with ribosomes

Viruses, by far, are the most notorious infectious agents that replicate only inside a living cell and can infect all types of life forms. Of these, retroviridae is a family of enveloped viruses that stores its genetic information in the form of RNA. Upon infection in the host cell, this RNA is reverse-transcribed and incorporated into host DNA, using viral and host proteins. Thereafter, the viral genes are expressed using the canonical transcription and translation pathways.

Human immunodeficiency virus (HIV) is a retrovirus that can lead to acquired immunodeficiency syndrome (AIDS), a condition in which the immune system relinquish to life-threatening opportunistic infections. It has a very high genetic variability resulting from a fast replication cycle combined with a high mutation rate. Next, its core genome is conserved with two copies of positive single stranded RNA, each carrying three ORFs encoding polyproteins- gag, pol and env. Gag (group specific antigen) comprises about 50% of the mass of viral particle and is the major structural protein of HIV-1. *Pol* encodes precursors of reverse transcriptase, integrase and protease, while *env* codes for surface glycoproteins which participate in fusion of virus into the host cell.

Gag is transcribed as a 9kb long unspliced transcript in the host cell, which upon translation yields the polyprotein containing four major domains; the N terminus matrix (MA), capsid (CA), nucleocapsid (NC) and the C terminus p6. Two "spacer" regions- p2 and p1 peptides- are present on either side of NC-p7 (Briggs et al., 2009). It has been shown that gag regulates the balance between RNA translation and its packaging into virions, in a bimodal manner. Gag initially acts in a positive feedback manner to stimulate its own production (Anderson and Lever, 2006). This is determined by the MA peptide interaction with various factors involved in translation, including eIF5B, EF1α but the mechanism is not clear (Cimarelli and Luban, 1999). Beyond a certain concentration of gag, translation is inhibited, dictated by the recognition of packaging signal, present on the viral RNA. This function of Gag could be mediated through the NCp7 region, which has been shown to interact with RNA as well as other cellular proteins (Anderson and Lever, 2006).

Table 4: Proteins known to interact with NC-p7 and their cellular functions.

NC-p7 Interacting partner	Cellular protein function	Viral RNA-dependent
Stau (staufen) (Chatel-Chaix et al., 2008)	It is a double-stranded RNA binding protein involved in RNA localization, mRNA decay. Also forms the principal component for transport of ribonucleoproteins.	No
Lyric (Lysine-rich carcino-embryonic antigen-related cell adhesion molecule co-isolated); astrocyte-elevated gene 1 (AEG-1) (Engeland et al., 2011)	It is a HIV inducible gene in astrocytes that promotes HIV replication. It has also been implicated in HIV associated neuropathy, antiapoptotic effects and tumorigenesis.	No
Kif4 (kinesin superfamily protein) (Kim et al., 1998)	It acts as a motor protein for transport toward cell membrane. It is abundant in juvenile neurons and lymphatic tissues.	Not known
ABCE1 (Lingappa et al., 2006)	It is highly conserved, ubiquitously present protein in archaea and eukaryotes. Required for ribosome biogenesis and ribosome recycling during translation.	No
TSG101 (human tumor susceptibility gene 101) (Garrus et al., 2001)	It is a component of the endosomal sorting complex ESCRT-I and is involved in sorting of ubiquitinylated cargo into small vesicles.	Not known
Alix (apoptosis linked gene 2 interacting protein X) (Popov et al., 2008)	It promotes viral egress and membrane fission events.	Not known
Nucleolin (Bacharach et al., 2000)	It shuttles between the nucleus and the cytoplasm and has an active role in ribosome biogenesis and assembly.	Not known

The 55 amino acid long NC protein has two zinc fingers separated by functionally important basic domain [406]RAPRKKG[412] (Godet et al., 2012). Both the zinc fingers are functionally distinct, with the first being important for RNA encapsidation and the second required for viral particle stabilisation. During viral assembly and maturation, the NC region is crucial for gag-gag interactions and during viral replication (Thomas and Gorelick, 2008).

NC has been researched extensively and a large range of its interacting partners have been reported (Table 4). Out of the seven interacting proteins with NC-p7, three affect translation by being involved in ribosome biogenesis, recycling or transport of mRNA for translation. Moreover, GAG polyprotein being involved in its own translation regulation (Anderson and Lever, 2006), has been observed to interact directly with most of the large and small subunit ribosomal proteins. However, the mechanism of interaction by far remains unclear and needs to be investigated in more detail.

1.21 Project Outline

At the time of beginning this project, the human ribosome structure had not been studied to high resolution. This signified a limitation for eukaryotic translation studies, like for mechanism of interaction of viral sequences, like IRES with ribosome; with regards to antibiotic side effects. Most antibiotics bind to the prokaryotic ribosomes but some of them interact with eukaryotic counterparts, leading to risky unpredictable outcomes. Human ribosome structure is a necessity to avert these situations. We took upon ourselves the task of studying the human ribosome and its structure-function aspects. In particular, to obtain a homogenous ribosome sample which can in turn, be used for numerous purposes, ranging from studying the human ribosome structure itself to forming functional complexes.

Specifically, we focussed on the eukaryotic translation termination, which seems to have diverged from prokaryotes to eukaryotes. The function of the release factors involved in translation is similar across the different domains of life, but their structures are considerably different. Moreover, the mechanism of stop codon recognition and coordination of the peptide release with stop codon recognition, have been a mystery. Also, in this regard the knowledge of high-resolution structures of eukaryotic release factors is inadequate. We started from exploring the structure of full-length eRF1-eRF3-GTP ternary complex and went on to provide insights into the interactions of these two release factors. This was followed by initial efforts to reconstitute translation termination complexes *in vitro*.

Thus, my thesis primarily focuses on the structure-function aspect of the two major unanswered questions in eukaryotic translation:

1. Human ribosome structure analysis. To work towards obtaining a high-resolution structure of human ribosome using a combination of state-of-the-art techniques in structural biology such as X-ray crystallography and cryo-electron microscopy (cryo-EM). It would provide insights into the specific mechanism of protein synthesis and regulation in humans as compared to other eukaryotes and bacteria, in particular.

2. Translation termination mechanism in mammals. An attempt was made to fill the gaps in the understanding of the molecular mechanism of translation termination in higher eukaryotes. This involved purification and co-crystallisation of the two

eukaryotic release factors (eRF1 and eRF3). Preliminary results shed light on the eRF1-eRF3 interactions using biochemical studies and X-ray crystallography. Majorly, the study focuses on cryo-EM analysis of the pre-termination complex assembled using the purified ribosomes and release factors.

Finally, another aspect that we tried to understand was the interaction of Gag, HIV-1 polyprotein with ribosomal components. This was pursued in collaboration with Yves Mély's group at the Faculty of Pharmacy, Illkirch. Proteins like Gag regulate their own translation by interacting with the translation machinery but the mechanism of these interactions is not yet revealed. We worked on the characterization of interactions between RPL7 and Gag and provide insights into the viral life cycle. These contacts can be exploited in future, for generating effective drugs against dreadful viral organisms.

2. Methods

2.1 Materials

All chemicals were purchased from Sigma-Aldrich and Fluka. Crystallization screening kits and PEG solutions were purchased from Hampton Research.

Deionized, distilled water is used for buffer preparations and complete protease inhibitor (Roche) is added to all the buffers. Buffer A contains 20 mM Tris pH 7.5, 2 mM Mg(OAc)$_2$ and 150 mM KCl. Buffer B contains 20 mM Tris pH 7.5, 6 mM Mg(OAc)$_2$, 150 mM KCl, 6.8% sucrose, 1mM DTT and RNasin Plus RNase Inhibitor (Promega). Resuspension buffer C contains 100 mM KCl, 5 mM Mg(OAc)$_2$, 20 mM HEPES pH 7.6, 1 mM DTT and 10 mM NH$_4$Cl. For 60S and 40S subunits purification a slightly modified buffer A is required containing 20 mM Tris pH 7.5, 2 mM Mg(OAc)$_2$, 500 mM KCl.

Required synthesized mRNA was purchased from Dharmacon, Thermo-fischer Scientific GmbH. tRNA-Lys (uncharged), purified from fresh chicken liver, was a kind gift received from Guillaume Bec, IBMC, Strasbourg.

2.2 Human ribosome purification

HeLa cell preparation

HeLa cells are grown in suspension cultures (55 X 10^8 cells, approx 6L) in Minimal Essential Media Spinner Modification (S-MEM) (Sigma Aldrich) supplemented with 7% newborn calf serum, 2 mM Glutamine and 40 µg/ml gentamycin at 37° C in 5% CO$_2$ environment. Once confluent, they are serum-starved for 6 hours to get a synchronized cell population. These cell cultures were maintained at the IGBMC HeLa cell facility.

Lysis and sucrose cushion

Cells are then lysed in freshly prepared lysis buffer containing 15 mM Tris pH 7.5, 0.5% NP40 (Sigma-Aldrich), 6 mM MgCl$_2$, 300 mM NaCl and RNasin (Promega). After 30 minutes incubation on ice, the lysate is centrifuged at 12000 X g for 10 minutes to remove debris, nuclei and mitochondria (Belin et al., 2010). The supernatant is loaded on 30% sucrose cushion prepared in Buffer A and centrifuged for 16 hours at 115800 X g (50.2 Ti rotor) to get the crude ribosomal pellet (Jan and Sarnow, 2002) (Matasova et al., 1991). While loading the cushion, care must be taken to avoid disturbing the sucrose cushion, and ensure slow

addition of the lysate to the 30% sucrose. This pellet is resuspended in Buffer B and mixed thoroughly to obtain a homogenous resuspension. The presence of non-resuspended particles in this solution can affect the next step, and these particles must be removed by a short centrifugation (10 minutes at 10000 X g). Only the supernatant is used for the next step.

Sucrose density gradient

Sucrose solution must be treated with bentonite after preparation with buffer A to inhibit ribonucleases (Jacoli et al., 1973; Tyulkina and Mankin, 1984) if present.

Gradient preparation: SG 50 Gradient Maker (GE Healthcare) is used to make a linear gradient of 15-30%, wherein the higher % sucrose solution is loaded in the mixing chamber and the lower % sucrose solution is loaded in the other, allowing to mix slowly. The outlet is connected with a pump and sucrose is collected drop-wise from the outlet.

The supernatant is treated with 1 mM puromycin for 30 minutes at 4°C (Blobel and Sabatini, 1971) with intermittent mixing and loaded on 15-30% sucrose gradient prepared in Buffer A. The samples are centrifuged at 25000 rpm for 11 hours in a SW-28 rotor and fractions are collected from bottom to top using an Econo Gradient Pump (Biorad) with an Econo UV Monitor (Biorad) and a Fraction Collector. The sample absorbance is recorded using UV reader (Biorad) and the peak corresponding to 80S is pooled for PEG20K precipitation (Ben-Shem et al., 2011). A final concentration of 7% PEG20K is added to the pooled fractions, incubated on ice for 10 minutes and centrifuged at 17400 X g for 10 minutes. The pure ribosomal pellet is dissolved in resuspension buffer C and filtered using 0.22 μm filters (Millipore) for further analysis or stored without filtration on ice for seven days for crystallization. Snap freezing and storage is not advised. The concentration of the 80S ribosome was estimated using the absorbance measured at 260nm, one absorbance (A_{260}) unit corresponds to 20 pmol of 80S ribosome.

Subunit preparation

After 1 mM puromycin treatment, the pellet in buffer B is loaded on 15-30% sucrose gradient, prepared in 20 mM Tris pH 7.5, 2 mM $Mg(OAc)_2$ and 500 mM KCl. Two peaks

corresponding to 60S and 40S are pooled separately, and thereafter used for 7% PEG20K precipitation as for 80S ribosomes.

1. SW-41 tubes and rotor were used for small scale preparations (500 ml to 2L of cell cultures)

2. Hela cells: pelleted at different stages growth phase, stationary phase; glutamine starvation for 1-2 hours, serum starvation for 2-6 hours.

3. Lysis: Using detergent NP-40, Triton X-100, homogeniser

4. Sucrose % for gradients: 10-30%, 15-30%, 5-30%.

5. Removal of sucrose after the gradient: dialysis using GeBaflex tubes- MWCO 6kDa, NAP-20 column, Zeba Spin 7K MWCO spin column, PEG precipitation and centrifuging the pooled fractions at high speed.

6. Resuspension buffer: without NH_4Cl, with KOAc in place of $Mg(OAc)_2$. Different concentrations of Magnesium were tested, ranging 2-20 mM.

2.3 Ribosome characterization

2.3.1 Analytical Ultracentrifugation (AUC)

Analytical Ultracentrifugation allows quantitative analysis of macromolecules in solution. As the sample is centrifuged, the components separated out into layers forming boundaries due to movement of particles (Kieft et al., 2007). Sedimentation is observed (measuring absorbance/fluorescence) in real time, used to calculate the Svedberg's coefficient (s) for the molecule. Sedimentation/ Svedberg's coefficient values, in turn, depend on the size, shape and interactions of macromolecules in solution.

$$s = \frac{v}{\omega^2 r}$$

Where s, sedimentation coefficient, v, terminal velocity, ω, is the angular velocity of the rotor and r is the distance of a particle to the rotor axis (radius). Using this and Lamm equation, it is possible to determine the hydrodynamic properties of a sample (Lebowitz et al., 2002).
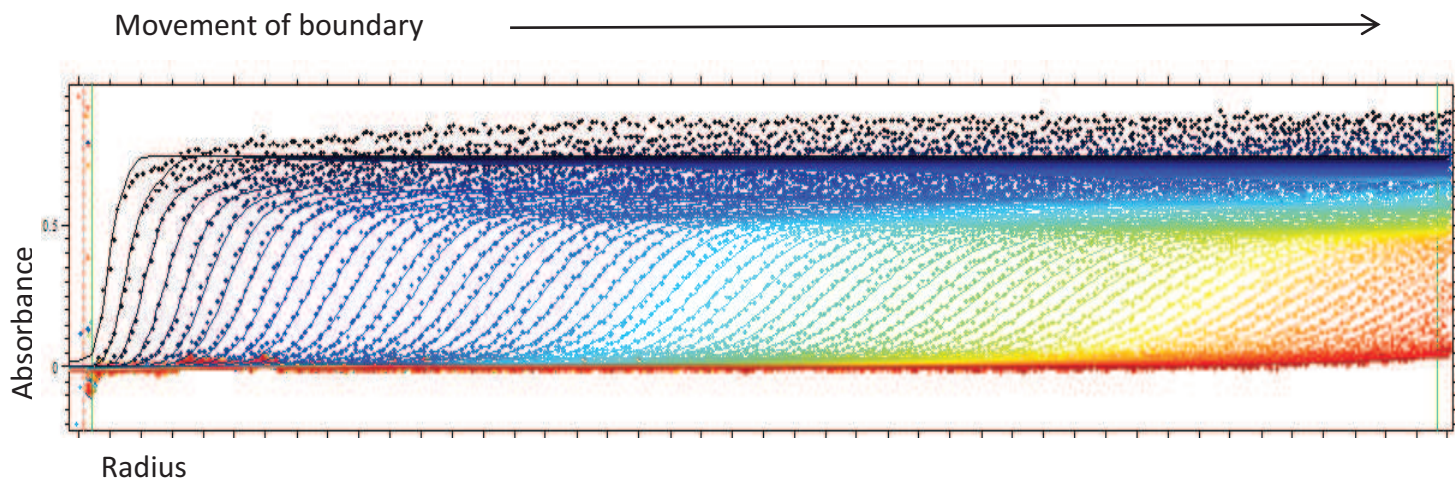
Figure 29: Data collection for velocity sedimentation. The data is collected as a set of boundaries as the 80S sediments during centrifugation. Each trace in this figure represents a boundary collected at 4 min intervals using UV-absorbance.

The Lamm equation can be written as:

$$\frac{\partial c}{\partial t} = D \left[\left(\frac{\partial^2 c}{\partial r^2} \right) + \frac{1}{r} \left(\frac{\partial c}{\partial r} \right) \right] - s\omega^2 \left[r \left(\frac{\partial c}{\partial r} \right) + 2c \right]$$

Where c is the solute concentration, *t* is time, *r* is radius, *D* is solute diffusion constant, *s* is sedimentation coefficient and *ω* is rotor angular velocity. On spinning the samples at constant angular velocity and recording the change in concentration *c(r, t)*, the parameters *s* and *D* can be determined.

Sedimentation velocity experiments were conducted using Beckman Coulter Proteome Lab XL-I analytical ultracentrifuge using the 8-hole Beckman An-50Ti rotor at 4°C for 80S, 60S and 40S samples in resuspension buffer. In fact, different preparations with changes in buffer composition were also analysed using AUC. Sedimentation at 15000 rpm was monitored by absorbance at 280 nm and 260 nm with scans made at 4 min intervals (Fig. 29). The solution density and viscosity for resuspension buffer were calculated using SEDNTERP software. Data was analysed using a c(s) model in SEDFIT.

2.3.2 Size Exclusion Chromatography Multi-angle Laser Light Scattering (SEC MALLS)

SEC column coupled with MALLS is a powerful technique to determine molecular weight and homogeneity of a sample. A SEC column is connected with a refractive index measuring

device and a lights scattering detector. When light passes through the fractions from the column, it is scattered depending on the interaction with the molecules. The intensity of light scattered is measured which determines the molar mass of protein.

80S ribosomal sample was analysed using Superdex 200 10/300 analytical column (GE healthcare life sciences) connected with Dawn DSP detector (Wyatt Technology, Santa Barbara, CA). To prevent bacterial growth, 0.01% sodium azide was added to the resuspension buffer, which was then filtered through 0.25 μm filter membranes (Millipore) prior to equilibrating the analytical column. The system was operated at 20°C, with a flow rate of 0.75 ml/min.

2.4 Single particle cryo electron microscopy (cryo-EM)

EM is a direct visualization method and can be used for structure determination of the sample. For cryo-EM, as the name suggests the sample is frozen in liquid ethane which allows formation of vitreous ice rather than crystalline ice to keep intact the structure of the macromolecule. This technique of freezing the sample to reduce radiation damage was first developed in 1981 (Dubochet et al., 1981). Initially, cryo-EM was used only for the study of highly ordered patterns, such as fibers of twisted protein (actin), two-dimensional crystals (like crystals having one or two lipid layers), and highly symmetrical structures, such as viruses. Since, only such kind of sample led to good resolution. 2D crystals were also studied over decades for membrane proteins such as (Kuhlbrandt, 1987; Kuhlbrandt and Downing, 1989; Unwin and Henderson, 1984) but these attempts at getting the ribosome structure failed (Yonath et al., 1987). The method of single particle 3D reconstruction came to the rescue where multiple copies of a single molecule in different orientation are used to reconstruct a 3D structure (Frank, 1990).

On irradiating a sample with electrons, the protein being denser than ice, appears as a dark area in comparison with background. But, in the frozen sample there is a problem of low contrast and it is critical in cryo-EM to see particles. There exist various methods to overcome this problem however; each has its own limitation. For example, increasing the dosage of electrons, improves the visibility but on the other hand, this causes radiation damage to the image. Another parameter that can affect visibility is the sample distance from the real focal point of the lens, i.e., defocus value. High defocus values, indeed

improve visibility, but it leads to loss of high frequency information, in turn, limiting the resolution of the final structure. Moreover, specimen drift, charging, radiation damage, ice thickness and contamination of the sample during data collection, beam drift and microscope stability problems, and newly discovered particle motion in the ice caused by electron beam are some of the problems, explained below, that can cause poor data quality. Thus, there needs to be a fine play between the various parameters to collect good data.

Firstly, in order to avoid beam damage to the sample, a relatively high dose of electrons is avoided. On the other hand, noisy images are recorded at low dosage exposures, where signal to noise ratio (SNR) remains below unity. Thus, there must exist a balance between electron exposures necessary to obtain useful SNR and electron dosage that avoids damage to the structural features. The tolerable levels of electron exposure have been determined to be 5-20 $\bar{e}/Å^2$. Thereby, a large number of copies of the molecule under "low dose" conditions are recorded which allows image alignment and then averaging to improve SNR and recover the high-resolution features.
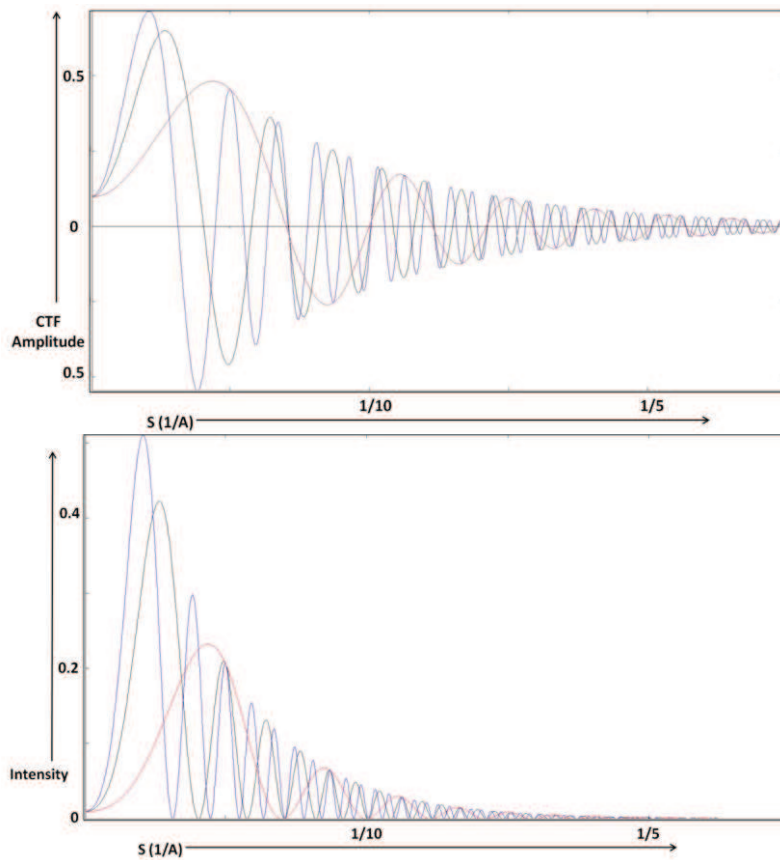
Figure 30: Top panel denotes the CTF amplitude at different defocus values. In red is -1 μm, black -2 μm and blue -5 μm defocus values. The lower panel shows the phase flipped CTF depiction.

88

Next, using different defocus value is the simplest way to vary contrast in microscopy. In the figure 30, the contrast transfer function (CTF), represented as a decaying sinusoid function, is shown at -1 μm (red curve), -3 μm (black curve) and -5 μm (blue curve) defocus values. The image with low defocus, close to real focus value -1 μm here, has its first peak of CTF shifted towards higher resolution but it has lower signal strength and thus, the problem of particle localization. Conversely, at -5 μm the first peak is at low resolution but has high intensity, thereby high visibility and contains global particle information. In the latter case, fine details about the particle are lost and this limits achievable resolution. Therefore, usually a defocus range is used (0.5 to 4 μm) to fill the missing information.

Moreover, the acceleration voltage needs to be chosen such that signal to noise ratio (SNR) versus resolution is optimum. At high voltage, the contrast tends to be lower due to less interaction with the specimen. Thus, small proteins (a few 100 kDa) can be better visualized at low voltage due to less inherent contrast while big molecules (with RNA/DNA) are usually imaged at high voltages (Reviewed in (van Heel et al., 2000)).

During data processing, the particles are selected semi-automatically from the images collected on the microscope and corrected for CTF which involves determination of the zeroes of the function for each collected image and flipping the negative part of the function to positive one. During reconstruction, all particles with different defocus are summed up and in order to avoid cancelling out the positive and negative values of the sinusoidal function, this operation termed as CTF-flipping, is necessary.

These particles are then centred by translational alignment and rotational alignment is used for determining relative orientations. The related particles with similar views are averaged to get classes with considerably higher SNR. The last step is 3D reconstruction based on the Euler angles to obtain the structure. This process of alignment, 2D classification, angular assignment and 3D reconstruction is carried out iteratively until a converged structure is obtained (Reviewed in (Orlova, 2000)).

In the past few years, several technical advancements have enhanced the attainable resolution by at least an order of magnitude for single particle ribosome studies. The advent of optically stable electron microscopes like Titan Krios and direct electron detection devices (DDD) have revolutionised the field. The insistent problem of image drift caused by manual

microscope handling is overcome by an enclosed platform combined with automatic sample loading in the Titan Krios microscope. This ensures thermal and mechanical stability of the sample during data collection.

Also, the new generation of detectors, the DDD have potential to improve signal transfer. CCD cameras used earlier have a scintillator to convert electron signals into photons which are then recorded into the CCD. This increases background noise by blurring the signal into an area of photons. Moreover, the high resolution data beyond 0.5-0.6 Nyquist frequency of detector tends to be down regulated. The complementary metal-oxide semiconductor (CMOS) cameras detect the incoming electrons directly, without the intermediate requirement of a scintillator. This allows attaining the Nyquist limit for 3D reconstructions and even beyond as observed in counting mode for Gatan K2 camera (Gatan Inc., Pleasanton, CA). These detectors have a particular property of making several read-outs during an exposure and can keep these so-called "sub-frames" before summing them to get a final image. Data processing using these "sub-frames", termed as "movie" processing, provides an effective way to compensate for previously described problems like sample drift or charging caused by the beam striking the sample, as it provides user the flexibility to choose desired "sub-frames". Furthermore, the problem of beam induced particle motion was not even noticed without "movie" processing approach.

In parallel, there have been new software developments like Relion (Scheres, 2012), based on maximum likelihood method. Its 2D classification approach can be used to simultaneously align and classify single-particle images in a reference-free fashion. In addition, the 3D classification in Relion is particularly good for samples with conformational heterogeneity since it sorts the data into subsets in an unsupervised manner, without prior knowledge about the structural variability present in the data. Thus, all these improvements together have provided an "order of resolution jump" in single particle cryo-EM specifically for ribosomes.

80S data collection

80S ribosome samples were deposited on Quantifoil 2/2 holey carbon film, blotted with a filter paper and flash-frozen (using FEI Vitrobot (Mark IV)) to obtain ribosomes embedded in a thin layer of vitreous ice suspended across the holes. The samples under different

preparation conditions were inspected visually to verify sample homogeneity. For the visually good samples with no aggregation and separate particle distribution on the grid, the images were collected. The data set was obtained at liquid-nitrogen temperature using the in-house FEI Tecnai F30 (Polara) FEG field emission gun transmission cryo electron microscope operating at 100kV or 300kV acceleration voltage or the in-house Titan Krios operating at 300kV acceleration voltage. Data collection parameters are given in Table 5 for two different sample preparations.

Table 5: Data collection parameters on F-30 Polara in-house electron microscope.

	Empty 80S ribosome	80S ribosome with E-site tRNA
Detector	CCD eagle 4K X 4K	CMOS falcon 4K X 4K
Voltage	100kV	300kV
Pixel size	1.82	1.14
Box size	240 X 240	512 X 512
Magnification	59k	93k
Number of particles	15,000	24,000
Total dose	15 ē/Å^2	20 ē/Å^2
Defocus	-0.8 to -3.5µm	-1 to -5 µm

Automatic image acquisition was performed using EPU software (FEI). After visual inspection only images with best power-spectra were selected for image processing. Particle selection was carried out semi-automatically in e2boxer.py (EMAN2) or gmpicker (in-house software) and subsequently, selected particles were visually inspected and validated. Defocus value estimation and contrast transfer function (CTF) correction by phase flipping were performed using the program e2ctf.py from the EMAN2 software package (Ludtke et al., 1999) (Tang et al., 2007). 3D reconstruction and refinement was carried out using the EMAN2 software packages (Ludtke et al., 1999) (Tang et al., 2007) and Relion (Scheres, 2012). The resolution of the final three-dimensional structures was estimated by Fourier Shell Correlation (Saxton and Baumeister, 1982) according to the 0.5, 0.14 (Rosenthal and Henderson, 2003) and one-half-bit (van Heel and Schatz, 2005) criteria. The final resolution obtained for empty ribosome reconstruction are to 13.3 Å, 8.7 Å, 9.3 Å, consistent with the features of the maps; while it was 16.5 Å, 11.5 Å, 11.47 Å for ribosome with E-site tRNA at 0.5, 0.14 and one-half-bit FSC, respectively.

2.5 Crystallization of human 80S ribosomes

Obtaining crystals of protein of interest is one of the critical steps in the high-resolution structure determination using X-ray crystallography. The protein of interest is maintained in a dissolved state in suitable buffer that keeps the natural state of the proteins. Introducing, suitable precipitating agent to the solvent can bring protein to its super saturated state. Prior to the supersaturated state, it reaches saturated state at which the crystal nuclei appear. When the protein reaches the super saturation state, it results in the formation of three dimensional crystals from the nucleation centre. For crystallization of macromolecules various diffusion methods are available, like vapour diffusion (sitting drop and hanging drop), counter diffusion (capillary) and batch method (oil drop). In the case of vapour diffusion methods, the mixture of protein and the precipitating agent is allowed to equilibrate against the precipitating agent in a sealed reservoir chamber. Initially, the protein is in solution at point "A", in the phase diagram (Fig. 31), as the drop attains equilibrium, it reaches the nucleation zone in the phase diagram where crystals appear. Thereafter, it falls back into metastable region which is most suitable for crystal growth, point "C". For the counter diffusion method crystals are grown in capillaries, where the kinetics is completely different. The diffusion occurs in a restricted geometry that results in a controlled diffusion mechanism. This minimizes the super-saturation, leading to a continuous gradient of super saturation phases in the reaction mixture. The nucleation front progresses forming a gradient of crystal sizes, from a shower of small crystal at the precipitant loading end to slowly increasing bigger crystals at the protein loading end (Ng et al., 2003).

Practically, this requires searching for the potential precipitant conditions where protein can reach supersaturation (or solubility minima) to form the critical nuclei (McPherson and Gavira, 2014). This can be achieved by altering few critical variants:

- Protein concentration and pH change (pH change affects the ionization state of surface amino acids) Altering the protein (pH change which affects the ionization state of surface amino acids)
- Equilibration of drop based on salting-in or salting out phenomenon
- Interactions between protein and solvent (addition of polymers/ ions)

- Temperature
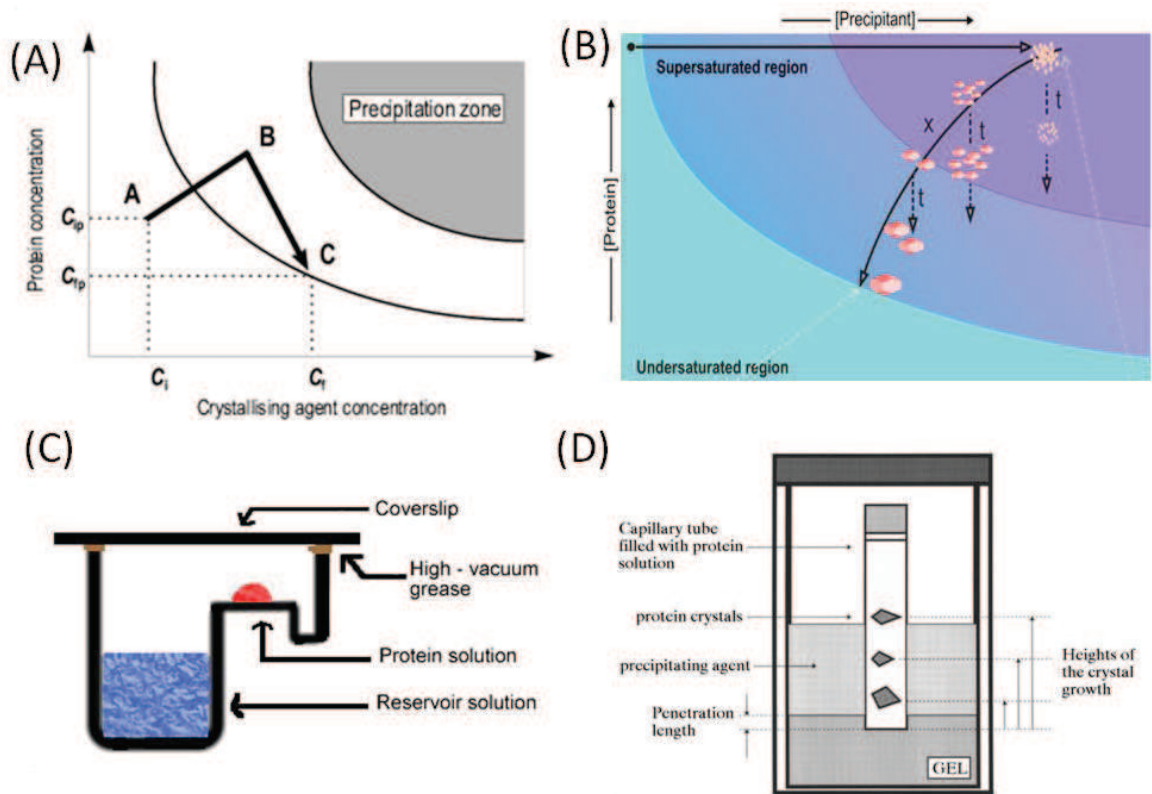
Figure 31: Comparison between capillary counter diffusion and conventional vapour diffusion. Phase diagrams and crystallization set up for sitting drop (A), (C) and capillary (B), (D). Adapted from Otálora et al., 2009.

Optimisation of the potential precipitant to achieve diffraction quality crystal is equally critical. The optimization process involves modulation of the rate of crystallization, seeding and finding one or more additives to promote crystallization. Seeding is a powerful technique to promote crystal growth. In this method, previously nucleated crystals are introduced in to the fresh drop of protein-precipitant mixture and equilibrated, in which these crystals act as nucleation centre for the protein of interest (Thaller et al., 1985). The additives are small molecules which can have drastic effects on macromolecular crystallization (additive screen- Hampton Research). Some additive kits act as a electrostatic cross linking agent that promote crystallization. In addition some additives prevent aggregation which includes detergents, solubilising agents, osmolytes and poisons (to reduce twinning) like DMSO, acetone etc.

These optimised crystals are diffracted using X-rays, which on hitting the crystal produce a characteristic diffraction pattern. The wavelength of X-rays, λ, is of the same order of

magnitude as the interplanar distance, d, in a crystal lattice which makes them ideal for diffraction.
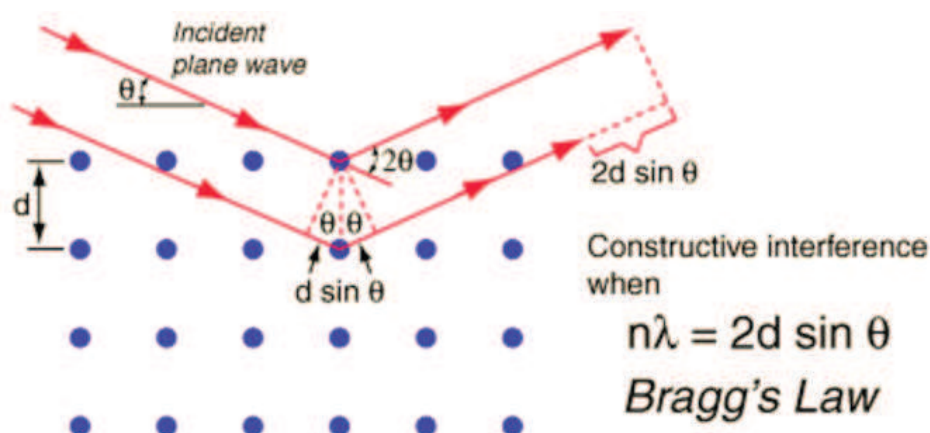
Figure 32: Bragg's law. (Thornton & Rex Sec. 6.1).

The Bragg's law describes the clause for constructive interference from two consecutive planes of crystallographic lattice. $2d\sin\theta$ is the path difference between two waves undergoing constructive interference, where θ represents the scattering angle. When scattered waves, satisfy the Bragg's law, very strong intensity spots, called Braggs peaks, are recorded on the detector. This diffraction pattern is the key in obtaining atomic crystal structure.

For crystallisation of 80S ribosomes, the purified 80S ribosomes were filtered using Millipore 0.22 µm after 7 days of annealing on ice (as for yeast ribosomes; (Ben-Shem et al., 2011), and the sample was kept at room temperature for an hour before setting up crystallization screens. To exploit the counter diffusion mechanism, 0.35 µl of purified ribosomes at 5-9 mg/ml were loaded on one side of The Crystal Former (Microlytic) and 0.35 µl of precipitant were loaded on the other side. A number of screens were tested including PEGs (Hampton Research), PEG-Ion pH (Hampton Research), Nucleix (Qiagen), and Protein Complex (Qiagen) at 4°C and 17°C. Microcrystals were obtained with 20% PEG 10K, 100 mM Na HEPES pH 7.5, or with 15% PEG 20K, 100 mM Na HEPES pH 7.5 at 17°C. Optimisation of the crystallization conditions with 30% PEG 20K, 10 mM Mg(OAc)$_2$ and 100 mM Na-HEPES pH 7.5 resulted in slightly larger crystals. Apart from the crystal former, also the crystal Harp (Molecular Dimensions) and classical glass capillaries of 0.5 mm and 1 mm were tested. In the process of optimization to get diffraction quality crystals, crystallization condition optimized in

counter diffusion was also used in vapour diffusion methods. Since the kinetics in these two diffusion methods are considerably different, the same crystallization reagents failed to provide crystals using vapour diffusion methods. Therefore, in a concentration range of 5-9 mg/ml of purified ribosomes, PEG 20K concentration and other conditions with PEG 10K as precipitant were optimized to obtain crystals in sitting drops. This included commercial screens such as PEGs (Hampton Research), PEG-Ion pH (Hampton Research), Nucleix (Qiagen) and Protein Complex (Qiagen). PEG 20K was varied from 0-20%, along with Magnesium 0-10mM. As concluded from drops, initially PEG 20K helped in solubilisation of ribosome, keeping the drops clear, but above 7-8%, the drops started to show increasing levels of precipitation. This helped us to narrow the search between 3 and 12%, which is just the border line for going from clear to precipitated drops. Plate-like crystals were obtained with 2 µl of 8 mg/ml of sample mixed with 2 µl of 4% PEG 20K, 100 mM Na-HEPES pH 7.5 and 50 mM KSCN. The crystals were tested with various cryo protectants which include, 18% glycerol alone, a series of small PEGs such as 8% PEG 4K, 8% PEG 6K, 8% PEG 8K along with 8% glycerol. Final concentration of 18% glycerol prepared in the above crystallization reagent resulted in better diffraction.

2.6 Release factors (eRF1 and eRF3) purification

2.6.1 Cloning

eRF1

The human eRF1 sequence containing clone was ordered from Invitrogen in pENTR 221, compatible with the Gateway cloning strategy. The LR reaction allowed to directly clone the sequence into pCoGWA (with N-terminal His tag, followed by a TEV protease site) for expression and protein purification. The clones obtained were screened using colony PCR and the sequence for the positive clone was confirmed by GATC® biotech sequencing.

eRF3

The human eRF3 gene (with a single base change of T to C, at position 1644) was synthesized *de novo* by GATC® biotech in pUC57 plasmid. Thereafter, the plasmid pUC57 was subjected to double digestion with restriction enzymes NdeI and BglII, alongside; the expression plasmid pnEAvHX was digested with NdeI and BamHI. The desired digested sequences were excised from agarose gel stained with Ethidium Bromide (EtBr), and ligated using T4 DNA Ligase of BioLabs for an hour. The ligated products were transformed into DH5α competent cells and the clones obtained were screened using single and double digestion. The positive clones were confirmed by GATC® biotech sequencing for the presence of the full eRF3 sequence.

2.6.2 Protein expression

Both eRF1 and eRF3 clones were transformed in BL21(DE3) and BL21(RIL) competent cells. One colony was picked and inoculated overnight in Luria Bertani (LB) broth containing the antibiotic ampicillin (100μg/ml). The expression of eRF1 and eRF3 were tested in LB, 2X LB and autoinducible media. BL21(DE3) worked well for eRF1 and BL21(RIL) for eRF3 expression. For protein expression and purification of eRF1, secondary culture was grown in 2X LB at 37°C, with IPTG induction at an O.D. (λ=600 nm) of 0.5-0.7 units, and incubated overnight at 16°C with continuous shaking. In contrast, for eRF3, auto inducible media was used; cultures were grown overnight at 37°C. After 16-18 hours, cells were pelleted at 4000 X g for 20 minutes. The conditions were optimized to get better yield, the cells were grown in 20L bio-fermentor.

2.6.3 Protein purification

The purification of proteins of interest was performed using a series of column chromatography system "AKTA purification" of GE Healthcare®. All these steps are performed at 4°C. At the end of every step, either 10% SDS PAGE gel or Labchip GX/GXII Caliper® was used to verify for the expression and purity of the sample.

Table 6: Buffer solutions used for every step of purification of eRF1 and eRF3. The role of glycerol is further explained in the text. PI is the protease inhibitor cocktail from Roche.

Step	eRF1	eRF3
1. Lysis	50 mM Tris HCl pH7.5, 200 mM KCl, 5 mM βME, PI.	50 mM Tris HCl pH7.5, 300 mM KCl 300, 10% glycerol, 5 mM βME 5, PI.
2. Ni affinity	Buffer A: 50 mM Tris HCl pH7.5, 200 mM KCl , 5 mM βME, PI. Buffer B: Buffer A + 1 M Imidazole.	Buffer A: 50 mM Tris HCl pH7.5 300 mM KCl, 10% glycerol, 5 mM βME, PI. Buffer B: Buffer A + 1 M Imidazole.
3. Dialysis	50 mM Tris HCl pH7.5, 50 mM KCl, 5 mM βME, 5% glycerol.	50 mM Hepes KOH pH7.5, 50 mM KCl, 10%glycerol, 5 mM βME.
4. Ion Exchange chromatography	Buffer A: 50 mM Tris HCl pH7.5, 10 mM βME, PI. Buffer B: Buffer A +1M KCl.	Buffer A: 50 mM Hepes KOH pH7.5, 10 % glycerol 10%, 10 mM βME; PI. Buffer B: Buffer A + 1M KCl.
5. Gel Filtration S75 16/60	50 mM Tris HCl pH 7.5, 300 mM KCl, 7 mM DTT , PI.	None.
6. Gel Filtration S200 for eRF1-eRF3 complex	50 mM Tris HCl pH7.5, 200 mM KCl, 10 % **glycerol**, 7 mM DTT, PI.	

eRF1 purification was performed based on the methods followed by Song and his co-workers (Song et al., 2000), while eRF3 purification required standardisation.

The cell pellets were resuspended using appropriate buffer, 0.001 g lysozyme Sigma® per gram of cells and DNaseI "RNase Free" Roche® (24U/g). The sample was lysed using sonication and followed by centrifugation at 40000 X g for 1 h at 4°C. The supernatant containing the soluble protein was filtered (0.45 μm) and loaded on the Ni affinity column. The fractions corresponding to the protein of interest were pooled and dialysed overnight against dialysis buffer, using a membrane with 3.5 kDa MWCO, at 4°C. This sample was loaded onto the ion exchange column. The chromatogram peak corresponding to the protein was pooled and concentrated using Amicon filter with 30 kDa MWCO (Millipore®). For eRF3, no further purification step was required before forming the complex. However, eRF1 went through an additional S75 gel filtration step. Subsequently, the pure protein was concentrated approximately to 10 mg/ml using Amicon filter with 30 kDa MWCO (Millipore®). Both eRF1 and eRF3 were either frozen in liquid nitrogen or used directly to form the complex.

2.6.4 eRF1-eRF3 complex purification

Purified eRF1 and eRF3 carry a His tag and His-thioredoxin tag, respectively at their N-terminus, separated by TEV protease recognition site (ENLYFQ(G/S)). Hence, to form the complex, eRF1 and eRF3 were mixed in 1:1 molar ratio, in the presence of TEV protease, added in 1:40 ratio. The sample was incubated overnight at 4°C and then loaded onto "Superdex 200 16/60" column. The purified protein complex was concentrated to 10-16 mg/ml and used for biophysical characterization and crystallization assays.

2.7 Biophysical characterization of purified release factors (eRF1 and eRF3) and eRF1-eRF3 complex

Mass spectrometry (MS)

MS allows measuring the molecular weight of the purified protein and all the other components or impurities present. eRF1, eRF3 and eRF1-eRF3 complex were analysed by Electron Spray Ionisation with a Time of Flight detector (ESI-TOF). 20 μL of minimum 50 μM concentration sample was dialysed overnight against a solution of ammonium acetate and used for ESI-TOF detection. The spectrum obtained was analysed for the expected molecular weights.

Dynamic Light Scattering (DLS)

To check the homogeneity, stability and polydispersity of the purified protein and protein complex, a DLS assaywas performed. It is based on the principle of laser light scattering by the particles in solution. The scattering intensity which is recorded fluctuates over time due to the Brownian motion of the particles, thereby giving information about the dynamic particles. For eRF1-eRF3 complex, 4μL of sample was placed in a micro quartz cuvette. The unit "DynaProNanostare" (Wyatt technology) emits a red laser of wavelength λ = 830 nm which passes through the sample across a length of 1 cm. The measuring beam provides information on the diffusion coefficient, hydrodynamic radius (R_d) and the polydispersity of the sample analysed.

Native PAGE

The native PAGE is used to determine proteins in their original state without denaturation, and is useful to detect protein-protein or protein-nucleic-acid interactions. 6%, 8% and 10% gels were used to examine the migration of eRF1-eRF3 complex with respect to eRF1 and eRF3 separately. The samples were loaded onto the polyacrylamide gel, placed in a Bio-Rad chamber for gels and ran with constant power of 2W for 2 hours at 4°C in migration buffer.

Microscale Thermophoresis Technology (MST)

MST is used to detect protein interactions. It is based on a microscopic temperature gradient that is generated by a laser beam. Any change of the hydration shell of biomolecules due to interaction with a binding partner during this temperature gradient is detected and is used to determine binding affinities and kinetics of interactions of the two components involved. The concentration of unlabelled protein (eRF1) was serial diluted into 16 tubes starting with 140 μM or the highest concentration achieved (360 μM). eRF3 was labelled at either cysteine or lysine residues with FITC labelling kit from Nanotemper and mixed at a constant concentration of 32 nM. The concentration of the fluorescently tagged protein to be used was adjusted by checking the signal for a series of dilutions. The 16 solutions containing serial diluted concentrations of eRF1 and constant eRF3 concentration were loaded in capillaries and placed in Nanotemper. Fluorescence measurements were made against a time scale and a graph was plotted for thermophoresis v/s concentration of unlabelled protein to detect the K_d. Both standard and hydrophilic capillaries were tested, with and without the presence of additional 0.05% Tween-20.

2.8 Crystallization of eRF1-eRF3 protein complex

Pre-crystallization tests (Hampton Research) were performed to determine the right protein concentration to use for crystallization assays. It prevents having most of the screen with precipitates or aggregates, and thus find a balance between precipitation and solubility to allow crystallization according to the phase diagram. 9-14 mg/ml of eRF1-eRF3 complex was mixed in 1:1 ratio with precipitant (100 + 100 nl or 200 + 200 nl drops), set up in MRC 2 drop plates and stored at 17°C and 4°C. The plates stored at 17°C were screened automatically in the FORMULATRIX storage but those at 4°C were screened manually. All the 16 commercial screens available at the IGBMC structural biology platform, total of 1536 conditions, were tested (Table 7). The initial hits obtained at 17°C were repeated in manual hanging and sitting drops (1.5 µl + 1.5 µl). In addition, the obtained crystallization condition was optimized with 1% and 2% additives (Hampton Research). Once optimized, GMPPCP was added to the protein complex and crystallization was set up using this sample.

Table 7: Commercial screens tested for eRF1-eRF3 protein complex.

Manufacturer	Hampton Research	Qiagen	Emerald Biosciences	Molecular Dimensions
Crystallization screen	Index SaltRx PEG/Ion	The Anions The Cations The AmSO4 The MDP The Classics The JSG+ The ProComplex The PEGs The Nucleix	Wizard I & II	Midas LBD

Also, the gel filtration buffer was modified to remove the 10% glycerol to check for sample stability. Since the DLS results for sample with and without glycerol were similar, the complex seems to be stable in both conditions. The 16 commercial screens were set up for crystallization with 9-16 mg/ml of protein complex. The crystals of eRF1-eRF3 complex in the presence of 10% glycerol diffracted to 8 Å. On the other hand, the crystals produced with the sample in the absence of 10% glycerol diffracted to 4 Å. Data collection statistics are given in the table 11 in section 3.2.4.

2.8.1 Cryoprotectant optimisation for the crystals

The choice of the right cryo-protectant solution is important to cause least damage to the crystal while freezing. In addition, cryoprotectants prevent ice formation of large amounts of water in the crystal during cooling, instead it forms amorphous, vitreous or glassy layer, which provides better diffraction. Several cryo-protectants were tested in accord with the original conditions. 20% glycerol, perfluoropolyethylene glycol (PF-PEG), paraffin oil and UCP (Ultimate cryoprotectant: 8% glycerol, 8% Ethylene glycol, 9% Sucrose, 2% Glucose) are some of the compounds which were used as cryoprotectants. The crystals were harvested using loops (Hampton Research) of variable size (40 microns to 500 microns in diameter) and frozen directly in liquid nitrogen stream at -160 °C. Also, for condition with 2M $(NH_4)_2SO_4$, no cryoprotectant was added, and the crystals were frozen directly from the drop. While for condition with Li_2SO_4, the precipitant was replaced with increasing concentrations of $(NH4)_2SO_4$ in the reservoir and drop was equilibrated overnight with the new reservoir (keeping all other components same in the mother liquor).

2.8.2 Crystal Data collection and processing

All the crystals were checked for X-ray diffraction at the PXII beamline at SLS (Villigen, Switzerland), with a wavelength 0.99 Å. Data was collected for the condition F with Li_2SO_4 as precipitant after grid-scan optimisation. The high-sensitivity Pilatus detector allowed data collection for weak diffracting crystals as well. Diffraction data was indexed using MOSFLM (Battye et al., 2011) to obtain the probable cell parameters and space group. The data was scaled and integrated using SCALA (CCP4 suite of programs (Winn et al., 2011)) (Evans, 2006) with particular attention to values of R_{merge} and I/σ. The phases were determined using molecular replacement in phaser (McCoy et al., 2007) with known eRF1 and eRF3 structures (PDB code 3E1Y, 1R5B, 1DT9).

The rigid body refinement of this structure was performed using Buster (Bricogne & Irwin, 1996), keeping a check on R_{work} and R_{free} values. Refinement is still ongoing as R_{free} values have not yet stabilised. Also, Contact utility in (CCP4 suite of programs) (Winn et al., 2011) was used to determine interacting residues with 5Å distance. And to calculate the solvent accessible surface area, naccess (Hubbard, S.J. & Thornton, J.M., 1993) was used. Figures were prepared using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.0.1 Schrödinger, LLC) and UCSF Chimera (Pettersen et al., 2004).

2.9 Post termination complex

The 80S ribosomes characterized during this work were used to study eukaryotic translation termination aspects. Required mRNA was synthesised by Dharmacon. Initially, a short sequence of 13 nt **DY547- GG ACC AAA UAA GG**, fluorescently labelled at the 5'-end was used to perform MST studies and study formation of the complex. Attempts were made to assemble the post termination complex using the above mentioned components. But probably the length of the mRNA was not long enough to form a stable complex with the ribosome. Possibly, the presence of 5'-tag hindered the complex formation. Thus, the sequence was modified to comprise 40 nt:

UUGGAAGAGGAAAUCGGGG ACC AAA UAA GGGGGAAGAGAU

It was designed synthetically and checked for any secondary structure formation using RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). The **AAA** sequence is to allow tRNA Lys (UUU) binding. Also, we tried to incorporate Kozak consensus sequence, which allows initiation in eukaryotic translation.

$$(gcc)\ gcc\textbf{\textit{Rcc}}AUGG$$

AUG represents the start codon, R refers to purine (A/G) and the lower case bases are the most commonly occurring in a gene.

UAA is the stop codon required for eRF1 recognition. Rest of the sequence was designed randomly, keeping a tab on secondary structure formation, which can make the mRNA unavailable to bind the ribosome.

12 pmoles of 80S ribosomes were incubated with mRNA (13-mer labelled with DY 547 at 5' end or 39-mer) and tRNA-Lys (UUU) in a molar ratio of 1:1.5:2 for 30 minutes at room temperature. eRF1-eRF3 protein complex was added in 5: 1 molar ratio with respect to 80S ribosomes and the complex incubated for another 60 minutes at room temperature before freezing directly on the Quantifoil 2/2 holey carbon grid. Single particle cryo-EM data collection and analysis was done as explained earlier for human 80S ribosomes with both mRNA sequences. The data processing scheme is mentioned in Figure 33.

Micrograph analysis by visual inspection

↓

Semi-automated particle picking

↓

CTF estimation and phase flipping

↓

2D classification

↓

Structure using C1 start up in EMAN2/
using a reference model in Relion

↓

3D classification and splitting into classes

EMAN2

Relion
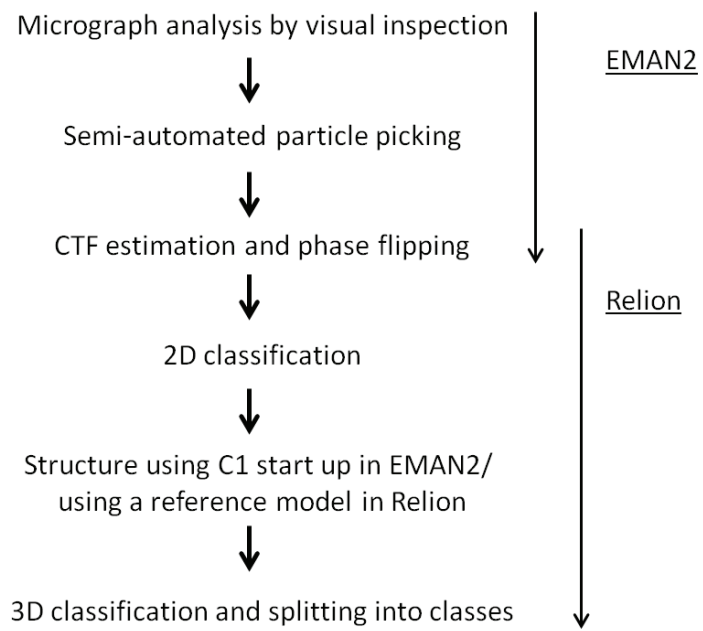
Figure 33: Pipeline for data processing in cryo-EM.

2.10 HIV-1 Gag interactions with ribosomal proteins

Our collaborators at the Faculty of Pharmacy, Illkirch found that interactions exist between Gag protein and RPL30 using yeast two-hybrid and pull-down assays. To validate these interactions with RPL30 bound to ribosome, we carried out the experiments described in the following section.

2.10.1 Polysome profile analysis

To study Gag protein interaction with ribosomes, sucrose density gradients (20-50%) were performed. The cells were lysed using buffer composed of 15 mM Tris pH 7.5, 6 mM $MgCl_2$, 25mM KCl, 0.5% Triton X-100, 2mM DTT, cycloheximide CHX 100μg/ml, RNAsin and protease inhibitor; and incubated for 10 minutes at 4°C. The extract was centrifuged for 10 minutes at 800 X g to remove cellular nuclei. The supernatant thus obtained was centrifuged again at 12500 X g for 10 minutes to remove mitochondria. Sucrose density gradients were prepared as mentioned before. The final supernatant was layered on 20-50% sucrose density gradients prepared in buffer A (20 mM Tris pH 7.5, 5 mM Mg(OAc)$_2$, 100 mM KCl) and centrifuged at 33000 rpm for 3 hours using a SW41 rotor. The gradients were fractionated from bottom to top. All the sucrose gradient fractions were precipitated with acetone to run SDS gels for western blot analysis. Ribosomal proteins and NCp7 were detected with polyclonal NCp7, S7e and monoclonal L24 antibodies.

In order to authenticate the specificity of these interactions, the cell lysate of gag transfected HeLa cells, after removing the nuclei and mitochondria, was treated with EDTA and run on sucrose density gradient. For these experiments, buffer A for sucrose density gradients did not contain magnesium. In addition, another experiment was performed using the cell lysate of gag transfected HeLa cells treated with puromycin and analysed for the density gradient profile.

Thereafter, negative controls for polysome profile were run with only EGFP (enhanced green fluorescent protein) transfected HeLa cells. In parallel, HeLa cells co-transfected with Gag and EGFP were lysed and analysed for their polysome profiles. EGFP being a reporter gene does not interact with ribosome, and thus its polysome profile would help rule out non-specific interactions.

2.10.2 Linear sucrose density gradient analysis NCp7

Using a set of deletion constructs, it was determined that NCp7 region of Gag was crucial for its interaction with RPL30. This was further verified by incubating the NCp7 peptide with human 80S ribosomes purified from HeLa cells (Khatter et al., 2014), for 30 minutes at 4 °C. Negative controls were samples with either 80S ribosome alone or NCp7 alone. These samples were loaded on 15-30% sucrose density gradients prepared in buffer A and centrifuged at 25000 rpm, for 10 hours and 30 min, in SW41 rotor. The gradients were fractionated and designated fractions corresponding to the peak of 80S were precipitated with acetone to run SDS gels for western blotting as done for polysome profile analysis.

3. RESULTS & CONCLUSIONS

3.1 Human 80S ribosome

The first phase of my project involved purifying 80S ribosomes to biochemical homogeneity which comprised setting up and optimising the protocol as described in the following article (Khatter et al., 2014). These purified ribosomes were characterized using an integrated approach encompassing various biophysical and structural biology tools.

Human ribosomes have been the focus of study for a long time (Anger et al., 2013; Boehringer et al., 2005; Matasova et al., 1991; Spahn et al., 2004b; Zenkova et al., 1991), but mostly they have been purified from human placenta or blood. I used HeLa cells as a source due to their abundant availability and ease of handling. The structures so far have remained elusive with respect to atomic details, the major restriction being the crystallisation process itself probably due to a large number of flexible rRNA regions.

The crucial aspect of this study was the use of cryo electron microscopy for screening the samples for absence of aggregation and homogeneity. We screened samples which had subtle changes in purification like different stages of cell culture, various magnesium concentrations and alternative methods to remove sucrose. This helped us to narrow down the best buffer conditions with 5 mM Mg and use this sample for crystallization. The first ever reported human ribosome crystals were obtained by counter-diffusion in capillaries. But due to lack of diffraction, it was necessary to crystallise in drops in the context of vapour diffusion. Several screens were made with different pH, salt and PEG concentrations, to study crystal growth in drops. The most interesting aspect was the effect of PEG 20K and Mg concentration. It was observed that appearance of precipitates in drops with respect to the PEG 20K and Mg concentration followed a bell-shaped curve. Lower concentrations of PEG 20K led to precipitation, while higher concentrations led to complete solubilisation and clear drops. Upon going from 2% to 24% PEG 20K, the drops showed precipitation until 6%, beyond this, the drops were clear. In the presence of 5 mM Mg, the transition from precipitation to clear was shifted forward to 10% PEG 20K. On increasing Mg concentration further to 10 mM, the PEG20K transition concentration further shifted to 12%. These results are in accord with the role of Mg for eukaryotic ribosomes where Mg concentrations higher than 7-10 mM lead to intermolecular interactions and that results in precipitation. We observed this Mg effect on ribosome distribution in cryo-EM as well, as the micrographs are clearly consistent with these biochemical observations (Fig. 34).
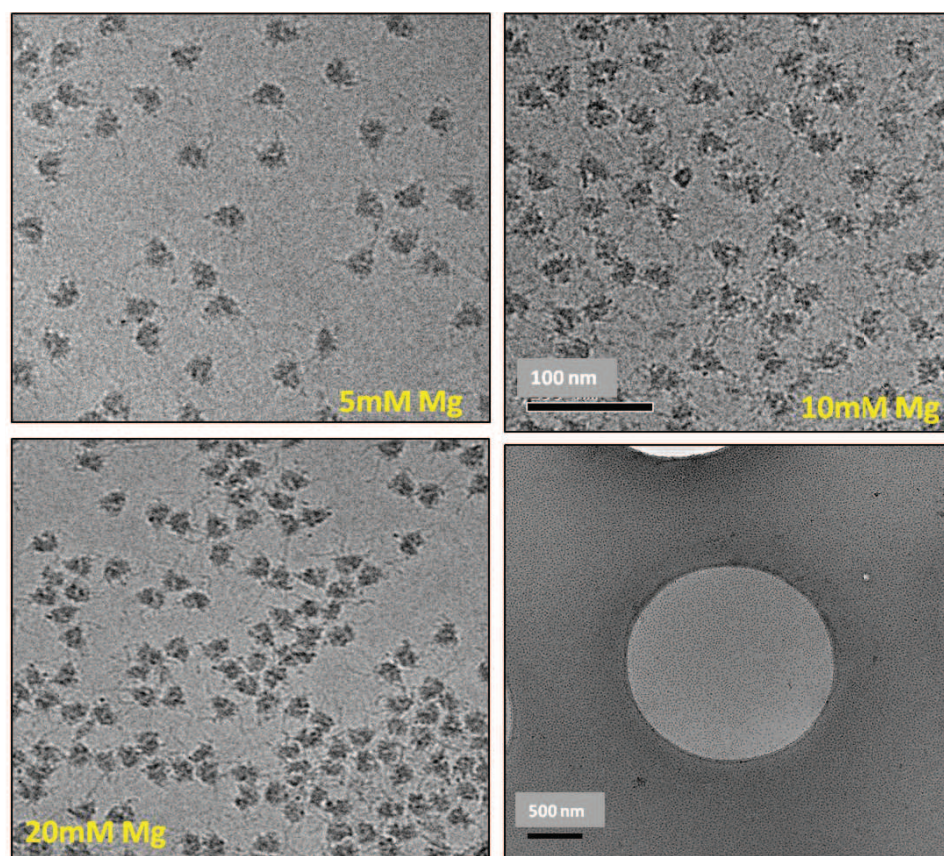
Figure 34: Cryo-EM images, depicting 80S particle distribution with different magnesium concentrations in resuspension buffer. The bottom right image is a low magnification view of the holey carbon grid showing a good distribution of particles. The scale bar denotes 100 nm and 500 nm.

This was followed by optimization in the zone of transition from precipitation to clear drops, 3-12% of PEG 20K. Subsequently, additives (Hampton Research) were screened with the suitable precipitant condition for crystals. Thin and fragile plate-like crystals with third dimension (2-10 µm) were obtained and the thickness of the crystals limits the inherent diffraction capacity of crystals. Although these crystals diffract up to 26 Å only, nevertheless they hold promise for future work.

Khatter, H., Myasnikov, A. G., Mastio, L., Billas, I. M. L., Birck, C., Stella, S., and Klaholz, B. P. (2014) **Purification, characterization and crystallization of the human 80S ribosome**, *Nucleic Acids Research* 1-11, 2014.

Purification, characterization and crystallization of the human 80S ribosome

Heena Khatter, Alexander G. Myasnikov, Leslie Mastio, Isabelle M. L. Billas, Catherine Birck, Stefano Stella and Bruno P. Klaholz*

Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), Centre National de la Recherche Scientifique (CNRS) UMR 7104/Institut National de la Santé de la Recherche Médicale (INSERM) U964/Université de Strasbourg, 1 rue Laurent Fries, 67404 Illkirch, France

ABSTRACT

Ribosomes are key macromolecular protein synthesis machineries in the cell. Human ribosomes have so far not been studied to atomic resolution because of their particularly complex structure as compared with other eukaryotic or prokaryotic ribosomes, and they are difficult to prepare to high homogeneity, which is a key requisite for high-resolution structural work. We established a purification protocol for human 80S ribosomes isolated from HeLa cells that allows obtaining large quantities of homogenous samples as characterized by biophysical methods using analytical ultracentrifugation and multiangle laser light scattering. Samples prepared under different conditions were characterized by direct single particle imaging using cryo electron microscopy, which helped optimizing the preparation protocol. From a small data set, a 3D reconstruction at subnanometric resolution was obtained showing all prominent structural features of the human ribosome, and revealing a salt concentration dependence of the presence of the exit site tRNA, which we show is critical for obtaining crystals. With these well-characterized samples first human 80S ribosome crystals were obtained from several crystallization conditions in capillaries and sitting drops, which diffract to 26 Å resolution at cryo temperatures and for which the crystallographic parameters were determined, paving the way for future high-resolution work.

INTRODUCTION

Ribosomes are composed of two subunits, the large (60S/50S) and the small (40S/30S) ribosomal subunits, which assemble together to form the functional 80S and 70S in eukaryotes and prokaryotes, respectively. Each subunit has protein and ribosomal RNA (rRNA) components with a relatively stable rRNA/protein ratio of 2:1 in cytosolic mammalian and bacterial ribosomes. The overall structure of the ribosome is conserved in all species consisting of the three tRNA binding sites [aminoacyl (A), peptidyl (P) and exit (E)], the GTPase center and the peptidyl transferase center. However, apart from the conserved core, eukaryotic ribosomes are more complex and contain many more proteins (26 extra) and longer rRNA (including long expansion segments, ES) (1). These ES have been hypothesized to allow ribosome docking on the endoplasmic reticulum, possibly providing scaffolding sites to bind additional proteins and form eukaryote-specific inter-subunit bridges (2). Also, functionally, the eukaryotic ribosomes have many more factors involved in every step of translation (initiation, elongation, termination and recycling) reflecting a high level of regulation (3–6).

The crystal structures of eukaryotic ribosomes from *Saccharomyces cerevisiae* 80S at 3.0 Å (7), *Tetrahymena thermophila* 60S at 3.5 Å and 40S at 3.9 Å (8,9) have highlighted the additional protein and rRNA components and precisely assigned their positions. The more complex higher eukaryotic ribosomes have been extensively studied by single particle cryo electron microscopy (cryo-EM) providing the first structure of the wheat germ ribosome obtained at 38 Å (10), which recently reached 5.5 Å resolution (11). Cryo-EM maps for mammalian

*To whom correspondence should be addressed. Tel: +33 388655755; Fax: +33 388653276; Email: klaholz@igbmc.fr
Present address:
Stefano Stella, Structural Biology and Biocomputing Programme, Centro Nacional de Investigaciones Oncológicas, C/Melchor Fernández Almagro, 3, E-28029 Madrid, Spain.

ribosomes such as canine, human (HeLa cell) and rabbit ribosomes (12–15) are available at relatively low resolutions of 8–15 Å but were the pioneer studies that allowed localizing eukaryote-specific proteins such as RACK1, rpS27e, rpS25e, rpL30e (16,17). Recently, the cryo-EM structures of human and drosophila ribosomes were elucidated in more detail with a resolution range of 5–9 Å, with the latest structure of the human ribosome reaching the 4–5 Å range (18). The availability of only a handful of eukaryotic ribosome structures (19–22) emphasizes the difficulty of appropriate purification, the limiting point being the availability of homogenous samples in large quantities required both for cryo-EM and crystallography.

Studying human ribosomes as opposed to ribosomes from other species is crucially important for understanding the mechanism of antibiotic action and selectivity with respect to ribosomes from various pathogenic bacteria, an ever increasing problem with the constantly growing occurrence of antibiotic resistance (23,24). The previous structures of prokaryotic ribosomes with antibiotics like paramomycin, streptomycin, tetracycline, hygromycin B, etc. have helped to elucidate the structural basis for their efficacy (25,26) and the species-specific interactions between ribosomes and antibiotics (27,28). Obtaining accurate information on the human ribosome would provide a prospect of developing specific antibiotics preferentially targeting the function of the prokaryotic ribosome with improved efficiency and reduced side effects, i.e. provide the molecular basis of cross-reactivity of existing or future antibiotics, which would be useful for the discovery of novel antibiotics. Human ribosomes have been purified earlier, from HeLa cells and placenta or blood (18,29) as separate subunits and reconstituted to 80S (30) for biochemical and cryo-EM analysis. However, to our knowledge, they have not yet been crystallized in the form of the fully assembled, endogenous 80S complex, which would be a key advance to get atomic level information. This would be crucial in particular for antibiotic complexes to understand the binding of those ligands that induce side effects, and of other and future drugs targeting the human ribosome (as accompanying drugs to reduce cellular activity such as required in the case of the treatment of cancer). Here, we establish a protocol to purify 80S ribosomes from HeLa cells in large amounts and describe the sample optimization by monitoring the homogeneity through sucrose gradients, Analytical Ultracentrifugation (AUC), Size Exclusion Chromatography Multiangle Laser Light Scattering (SEC-MALLS) and cryo-EM, which allowed obtaining crystallizable material. The crystals obtained here were characterized at synchrotron X-ray sources with respect to their crystallographic parameters such as cell parameters, space group, solvent content, etc., which represents essential information for future high-resolution work using crystallography.

MATERIALS AND METHODS

Detailed protocol of ribosome purification from HeLa cells

Equipments required: 10-l flasks for cell culture, SW-28 rotor, Type 50.2 Ti Beckman-Coulter rotor, GE SG-50 Gradient maker, Econo UV Monitor (Biorad), a Fraction Collector (Biorad), Econo Gradient Pump (Biorad).

Deionized distilled water is used for buffer preparations, and complete protease inhibitor (Roche) is added to all the buffers. Also, sucrose solution must be treated with bentonite after preparation with buffer A to inhibit ribonucleases (31,32) if present. Buffer A contains 20 mM Tris, pH 7.5, 2 mM $Mg(OAc)_2$, 150 mM KCl. Buffer B contains 20 mM Tris, pH 7.5, 6 mM $Mg(OAc)_2$, 150 mM KCl, 6.8% sucrose, 1 mM DTT, RNasin Plus RNase Inhibitor (Promega). Resuspension buffer C contains 100 mM KCl, 5 mM $Mg(OAc)_2$, 20 mM HEPES, pH 7.6, 1 mM DTT, 10 mM NH_4Cl. For 60S and 40S subunit purification a slightly modified buffer A is required containing 20 mM Tris, pH 7.5, 2 mM $Mg(OAc)_2$, 500 mM KCl. The role of ion concentration in inter- and intra-subunit interaction is discussed in the 'Results' section.

Step 1: HeLa cell preparation

HeLa cells are grown in suspension cultures (55×10^8 cells, ~6 l) in Minimal Essential Media Spinner Modification (S-MEM; Sigma Aldrich) supplemented with 7% newborn calf serum, 2 mM Glutamine and 40 µg/ml gentamycin at 37°C in 5% CO_2 environment. Once confluent, they are serum-starved for 6 h to get a synchronized cell population.

Step 2: Lysis and sucrose cushion

Cells are then lysed in freshly prepared lysis buffer containing 15 mM Tris, pH 7.5, 0.5% NP40 (Sigma-Aldrich), 6 mM $MgCl_2$, 300 mM NaCl, RNAsin (Promega). After 30 min incubation on ice, the lysate is centrifuged at 12 000g for 10 min to remove debris, nuclei and mitochondria (33). The supernatant is loaded on 30% sucrose cushion prepared in Buffer A and centrifuged for 16 h at 115 800g (50.2 Ti rotor) to get the crude ribosomal pellet (29,30). While loading the cushion, care must be taken to not disturb the sucrose, and ensure slow addition of the lysate to the 30% sucrose. This pellet is resuspended in Buffer B to homogeneity. The presence of nonresuspended particles in this solution can affect the next step, and these particles must be removed by a short centrifugation (10 min at 10 000g). Only the supernatant is used for the next step.

Step 3: Sucrose density gradient

Gradient preparation: SG 50 Gradient Maker (GE Healthcare) is used to make a linear gradient of 15–30%, wherein the higher % sucrose solution is loaded in the mixing chamber and the lower % sucrose solution is loaded in the other, allowing to mix slowly. The outlet is

connected with a pump and sucrose is collected drop wise from the outlet.

The supernatant is treated with 1 mM puromycin for 30 min at 4°C (34) with intermittent mixing and loaded on 15–30% sucrose gradient prepared in Buffer A. The samples are centrifuged at 25 000 rpm for 11 h in a SW-28 rotor and fractions are collected from bottom to top using an Econo Gradient Pump (Biorad) with an Econo UV Monitor (Biorad) and a Fraction Collector. The sample absorbance is recorded using UV reader (Biorad) and the peak corresponding to 80S is pooled for PEG20K precipitation (7). A final concentration of 7% PEG20K is added to the pooled fractions, incubated on ice for 10 min and centrifuged at 17 400*g* for 10 min. The pure ribosomal pellet is dissolved in resuspension buffer C and filtered using 0.22 µm filters (Millipore) for further analysis or stored without filtration on ice for 7 days for crystallization. Snap freezing and storage is not advised. For concentration calculations, 1 A_{260} unit corresponds to 20 pmol of 80S ribosome.

Analytical ultracentrifugation

Sedimentation velocity experiments were conducted using Beckman Coulter ProteomeLab XL-I analytical ultracentrifuge using the 8-hole Beckman An-50Ti rotor at 4°C for samples in resuspension buffer (35). Sedimentation at 15 000 rpm was monitored by absorbance at 280 nm with scans made at 4 min intervals. The solution density and viscosity for resuspension buffer were calculated using SEDNTERP software. Data were analyzed using a c(s) model in SEDFIT (Figure 1) (36).

Size exclusion chromatography multiangle laser light scattering

The molecular weight and homogeneity of the sample was checked using a SEC column coupled with MALLS Dawn DSP detector (Wyatt Technology, Santa Barbara, CA, USA) (Figure 1B). To prevent bacterial growth, 0.01% sodium azide was added to the resuspension buffer, which was then filtered through 0.25 µm filter membranes (Millipore) before equilibrating the Superdex 200 10/300 analytical column (GE healthcare life sciences). The system was operated at 20°C, with a flow rate of 0.75 ml/min.

Single-particle cryo-EM

For cryo-EM analysis, 80S ribosome samples were deposited on Quantifoil 2/2 holey carbon film, blotted with a filter paper and flash-frozen [using FEI Vitrobot (Mark IV)] to obtain ribosomes embedded in a thin layer of vitreous ice suspended across the holes (Figure 2). The images were collected at liquid-nitrogen temperature using the in-house FEI Tecnai F30 (Polara) field emission gun (FEG) transmission cryo electron

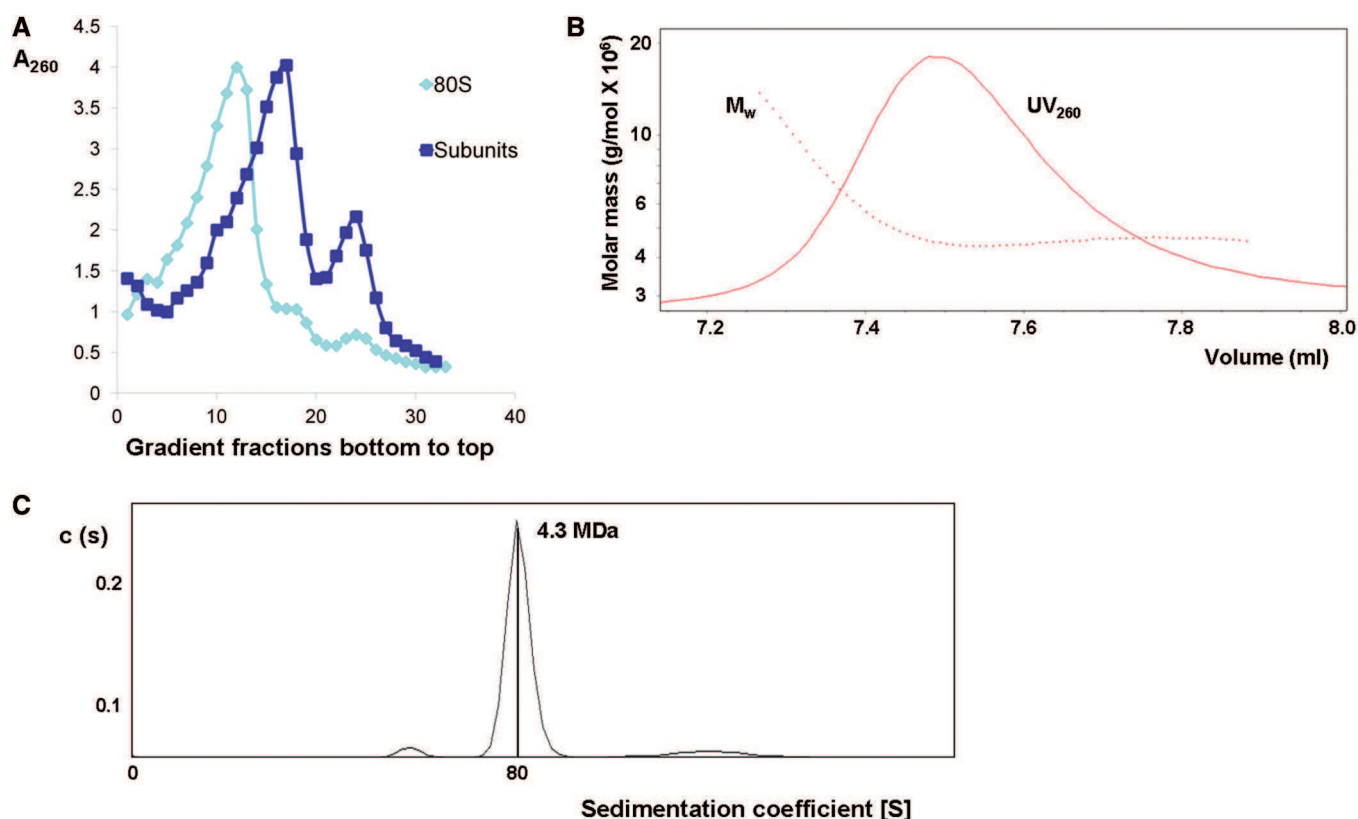**Figure 1.** Biophysical characterization of human 80S ribosomes purified from HeLa cells. (**A**) 15–30% sucrose density gradient profiles for 80S, and dissociated 60S and 40S subunits depicting the separation of components based on density. (**B**) SEC-MALLS for 80S sample and (**C**) AUC results showing the homogeneity of 80S sample, which corresponds to the calculated molecular weight of 4.3 MDa.

Table 1. Experimental details of cryo-EM data collection

cryo-EM data collection	Empty 80S ribosome	80S ribosome with E-site tRNA
Detector	CCD Eagle 4 k × 4 k	CMOS Falcon 1 4 k × 4 k
Voltage	100 kV	300 kV
Pixel size	1.82 Å	1.14 Å
Box size	240 × 240	512 × 512
Magnification	59 k	93 k
Number of particles	15 000	24 000
Total dose	15 $\bar{e}/Å^2$	20 $\bar{e}/Å^2$
Defocus	−0.8 to −3.5 μm	−1 to −5 μm

microscope operating at 150 kV acceleration voltage with a dose of ∼10–15 electrons per $Å^2$ at a magnification of 59 000 on a 4 k × 4 k CCD Eagle camera (FEI) resulting in a final step size of 1.82 Å per pixel. Another data set was collected at 150 kV acceleration voltage with magnification of 93 000 on a CMOS Falcon 1 camera (FEI; installed as an upgrade on the microscope while this work was ongoing) resulting in a final step size of 1.14 Å per pixel (Table 1). Automatic image acquisition was performed using EPU software (FEI). After visual inspection only images with best power-spectra were selected for image processing. Particle selection was done semi-automatically in e2boxer.py (EMAN2) with validation of all boxed particles by visual inspection. Defocus value estimation and contrast transfer function correction by phase flipping were done by using the program 2ctf.py from the EMAN2 software package (37,38). Structure determination and refinement was done using the EMAN2 software packages (37,38). The resolution of the final 3D structures was estimated by Fourier Shell Correlation (39) according to the 0.5, 0.14 (40) and one-half-bit (41) criteria to 13.3, 8.7 and 9.3 Å for the empty ribosome consistent with the features of the maps, while it was 16.5, 11.7 and 11.5 Å for ribosome with E-site tRNA, respectively.

Crystallization of human 80S ribosomes

The purified 80S ribosomes were filtered using Millipore 0.22 μm after 7 days of annealing on ice [as for yeast ribosomes; (7)], and the sample was kept at room temperature for an hour before setting up crystallization screens. Purified ribosomes (0.35 μl) at 5–9 mg/ml were loaded on one side of The Crystal Former (Microlytic) and 0.35 μl of precipitant were loaded on the other side. A number of screens were tested including PEGs (Hampton Research), PEG-Ion pH (Hampton Research), Nucleix (Qiagen) and Protein Complex (Qiagen) at 4 and 17°C. Microcrystals were obtained with 20% PEG 10 K, 100 mM Na HEPES, pH 7.5, or with 15% PEG 20K, 100 mM Na HEPES, pH 7.5, at 17°C (Figure 3A). Optimization of the crystallization conditions with 30% PEG 20 K, 10 mM $Mg(OAc)_2$ and 100 mM Na-HEPES, pH 7.5, allowed obtaining slightly larger crystals. Apart from the crystal former, also the crystal Harp (Molecular Dimensions) and classical glass capillaries of 0.5 and 1 mm were tested. These crystals could not be directly reproduced in sitting drops

(with 1:1 sample/precipitant) with the same conditions. Therefore, in a concentration range of 5–9 mg/ml of purified ribosomes, optimization of the PEG 20K concentration and other condition with PEG 10K as precipitant, was done to get crystals in sitting drops. This included commercial screens such as PEGs (Hampton Research), PEG-Ion pH (Hampton Research), Nucleix (Qiagen) and Protein Complex (Qiagen). Plate-like crystals were obtained with 2 μl of 8 mg/ml of sample mixed with 2 μl of 4% PEG 20 K, 100 mM Na-HEPES, pH 7.5, and 50 mM KSCN. Eighteen percent glycerol or a series of small PEGs such as 8% PEG 4K, 8% PEG 6K, 8% PEG 8K along with 8% glycerol as cryo-protectants were tested, with 18% glycerol alone being the best one found here.

RESULTS

Preparation and characterization of human 80S ribosomes

Because serum starvation is known to inhibit translation in eukaryotic cell lines (42), we reasoned that this may be a way to obtain more homogenous monosomes. HeLa cells were therefore grown in suspension balloons and serum-starved for 6 h or processed directly. Human 80S ribosomes were isolated from the cell lysate after centrifugation to remove mitochondria and nuclei, and the supernatant was loaded onto a 30% sucrose cushion. The pellet obtained was treated with puromycin to remove nascent peptides bound to the 80S, giving a yield of 7–8 mg of pure ribosomes from 6 l of cells ($20–25 × 10^8$) after the sucrose gradient (29). The presence of 18S and 28S rRNAs was confirmed by ethidium bromide staining. Several steps had to be optimized during purification to ensure homogeneity of the sample to be used for structural studies. Cell lysis being the most important step due to the presence of numerous proteases was treated with the maximum precaution. Several detergents such as CHAPS, NP-40 and Triton X-100 were tried, also dounce homogenizer was used to see if cell disruption would work better. However, 0.5% NP-40 turned out to be the mildest procedure, as monitored by cryo-EM imaging of 80S particles. The other bottleneck was removal of sucrose, which is a basic necessity for EM to ensure a good image contrast (43). For this, different methods were tried such as (i) centrifugation of the sucrose gradient fractions and resuspension of the pellet, (ii) exchanging buffer with a NAP column and then concentrating the sample and (iii) PEG 20K precipitation. The fastest, most convenient and least influencing procedure on ribosome stability was precipitation using PEG 20K, a method also used for yeast ribosome purification (7). Salt composition and concentration (KCl and Mg^{2+}) were optimized using cryo-EM imaging (see below).

The presence of all ribosomal proteins was confirmed using mass spectrometry (Supplementary Tables 1 and 2). It can, however, not be excluded that some loosely bound ribosomal proteins are present with variable stoichiometry in the 80S complex because protein peptide fragment occurrence depends also on how well the proteins can be digested and detected by mass spectrometry. The sample

was analyzed by SEC-MALLS and AUC (Figure 1B and C). The sedimentation coefficient was determined to be close to 80S with a single major peak obtained in AUC corresponding to 4.3 MDa (the theoretical value is 4.2 MDa), a result consistent with the chromatogram from SEC-MALLS. While the precision of SEC-MALLS and AUC cannot address whether all ribosomes contain all ribosomal proteins because of the high molecular weight, these analytical techniques were useful to monitor the sample quality throughout the purification procedure.

Cryo electron microscopy

To evaluate the homogeneity and monodispersity of the 80S ribosomes, the particle distribution was analyzed under different preparation conditions using cryo-EM imaging (Figure 2). This helped in finding the optimal conditions for purification, the optimal buffer being

Figure 2. Cryo electron micrograph depicting the distribution of 80S particles at 59k magnification, collected using a Polara F-30 electron microscope. The ribosomes were isolated from Hela cells (**A**) under normal growth conditions (**B**) in stationary phase of growth (**C**) after glutamine starvation (**D**) after serum-starvation (inset shows the zoomed-in view for individual ribosomes). (**E**) The cryo-EM structure of empty human 80S and with (**F**) E-site tRNA; both isolated from HeLa cells; under high and low KCl concentration during Puromycin treatment, respectively; color code: 40S golden, 60S blue, tRNA red.

150 mM KCl and 5 mM Mg^{2+} or less. Buffers with higher Mg^{2+} concentration showed aggregation, while higher KCl concentration increasingly leads to the appearance of the individual ribosomal subunits. The role of monovalent and divalent cations on subunit association has been well documented for prokaryotes and to some extent for eukaryotes as well. Previously, the concentration of potassium and magnesium in the resuspension buffer was observed to be crucial (44,45). In our hands, the sample completely aggregates at 20 mM $Mg(OAc)_2$, while at 5 mM and below it is monodisperse, as can be judged from the even distribution of ribosome particles on a cryo-EM grid (Figure 2C). This is consistent with the cellular concentration of magnesium being around 1 mM, which would promote the ribosomes to remain monodisperse when not translating. In contrast, the potassium concentration was found to be critical for the association of the two ribosomal subunits (46). At 150 mM KCl, 80S ribosomes were found to be stable at a concentration that compares well with that in cells. Only at significantly higher salt concentrations such as 500 mM KCl, the subunits become dissociated on the sucrose density gradient (Figure 1A). Consistent with this observation is the idea that potassium is required for inter-subunit stability while magnesium ensures proper rRNA folding. However, at higher concentrations, magnesium promotes inter-particle contacts, which lead to ribosome aggregation and precipitation as visualized by cryo-EM imaging.

Using optimized buffers, single-particle cryo-EM data sets were collected for 80S samples obtained from serum-starved HeLa cell for two salt concentrations during Puromycin treatment (150 or 300 mM KCl). The serum starvation is an important step towards getting a clean homogenous sample as was observed from direct cryo-EM imaging. Ribosomes isolated from nonstarved cells show impurities (Figure 2A), while ribosomes isolated from cells in stationary phase appear to undergo partial degradation (Figure 2B; for these reasons 3D reconstructions were not attempted from these samples). Glutamine starvation also shows ribosomes with either partial degradation or with impurities (Figure 2C). Serum is known to affect translation because its absence reduces the rate of *in vitro* polypeptide synthesis. We observed that on serum starvation of HeLa cells for 6 h, which leads to cell synchronization, ribosomes appear to be more homogenous, as can be judged from the cryo-EM images (Figure 2D).

A 3D reconstruction obtained from only 15 000 particles of 80S ribosomes (serum-starved) already provides a well-defined structure of the human ribosome with all the major landmarks such as the L1 stalk, the central protuberance in the large subunit and the head, beak, body and feet in the small subunit (Figure 2E). These features are consistent with the human ribosome structure published in previous studies (13,47). This 80S structure was obtained from ribosomes that went through a high-salt (300 mM) washing step during puromycin treatment right after the sucrose cushion. Next, to keep the salt concentration constant throughout the purification process, we decided to skip the high-salt washing step and maintain 150 mM KCl throughout. Interestingly, the cryo-EM reconstruction from such
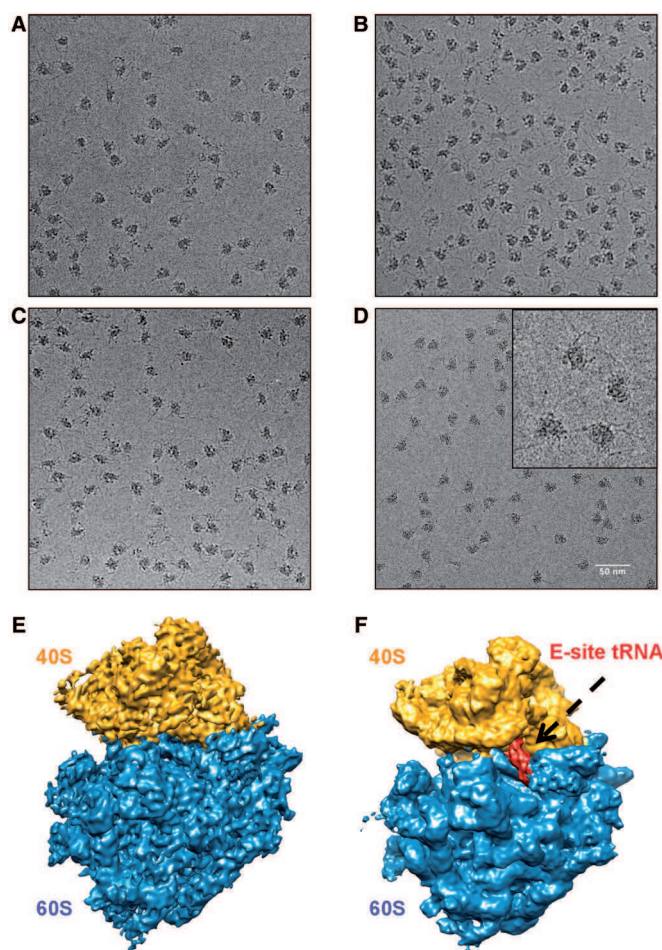
samples reveals the presence of an E-site tRNA (Figure 2F), a well-defined L1 stalk next to the E-site tRNA and well-ordered P-proteins (these regions are largely disordered in the high-salt washed ribosomes). This shows that the absence of a high-salt step allows maintaining an endogenous tRNA in the E-site, while mRNA and peptide are absent as revealed by the comparison of the corresponding cryo-EM reconstructions. The lower resolution of the E-site tRNA complex as compared with that of the empty ribosome may be due to ribosome subpopulations with different conformations, which would be compatible with the observation of different conformations in the asymmetric unit of yeast (7) and *T. thermophila* ribosomes (48).

Crystallization of the human 80S ribosome

The high homogeneity of the sample allowed us obtaining crystals in the microlytic crystal former capillaries as well as in a crystal harp, and also in sitting drops. After screening several conditions from PEGs (Hampton), PEG Ion pH (Hampton) and Protein complex (Qiagen) screens at 4 and 17°C, initial hits of microcrystals were obtained in capillaries with two conditions: 20% PEG 10K, 0.1 M Na HEPES, pH 7.5, and 15% PEG 20K, 0.1 M Na HEPES, pH 7.5, at 17°C. An increasing gradient of microcrystals (10–30 μm in size) could be observed from the precipitant loading side to the protein loading side. Optimization of the conditions provided larger crystals with 30% PEG 20K, 0.1 M Na HEPES, pH 7.5, and 10 mM Mg(OAc)$_2$ (Figure 3A). *In situ* diffraction from capillaries was attempted at the PX III beamline of the Swiss Light Source (SLS) synchrotron at room temperature as well as under cryo conditions (by freezing the tip of the capillary containing crystals, tested on beamline PX I). While no diffraction spots could be seen, this was a first indication that these crystals were not salt. However, owing to the complexity of handling them in capillaries they could not be mounted in loops. Direct freezing of capillaries was successful with respect to cryo conditions (no ice rings), but showed no diffraction (either no diffraction or high background due to the large mass of solvent in the capillary). Reproducing crystals under the same conditions in sitting drops was not successful probably because capillaries work on the principle of counter diffusion, which is strikingly different from vapour diffusion used in sitting or hanging drops in terms of the kinetics of equilibration and concentration used. Therefore, PEG 20K concentration was varied in a large range for forming crystals in sitting drops. Crystals with plate morphology were finally obtained in sitting drops (Figure 3B) with 4% PEG 20K, 100 mM Na HEPES, pH 7.5, and 50 mM KSCN in the reservoir. For crystal handling, mounting in loops (Figure 3C) and freezing, overnight stabilization with increasing PEG 20K in reservoir to 6 or 8% was tried, to stabilize crystals but resulted in no diffraction. Adding medium-sized PEGs such as PEG 4K, 6K or 8K with or without glycerol, either resulted in no diffraction or did not improve diffraction. First diffraction spots up to 60 Å could be observed for small, thin and fragile crystals (60 × 10 × 2 μm, SLS beamline PX I) with 16% glycerol as cryo-protectant added directly into the crystallization drop. With notably larger crystals (100 × 50 × 5 μm) diffraction improved to 26 Å (at SLS beamline PX I) showing a full reciprocal lattice (Figure 3E). For this and all subsequent experiments, the cryo-protectant was increased in steps of 2–4% to prevent osmotic shock damage to the crystal. The sample stability under crystallization conditions was verified using ethidium bromide–stained agarose gels for presence of rRNA after 2–4 weeks of sample preparation (Figure 3D). The gel shows 18S and 28S rRNA bands, as do ribosomes that were stored at −80°C after preparation. A full data set could be collected to ~40 Å resolution (Figure 3E and Table 2) from which the cell parameters, space group, Matthews coefficient, solvent content and number of molecules in the asymmetric unit could be derived (see below). Data were collected on a Pilatus detector at the SLS PX II beamline with a 10 × 30 μm or 10 × 50 μm X-ray beam and 0.2° oscillations, detector distance 1200 mm, wavelength 0.997 Å. Both 80S complexes with or without endogenous E-site tRNA, as characterized by cryo-EM (Figure 2), were used for crystallization assays. However, crystals grew only for ribosomes prepared at 150 mM KCl throughout, i.e., without any E-site tRNA washing step. This indicates that the crystals contain E-site tRNA, and that its presence could favor crystallization.

DISCUSSION

Human ribosomes represent an important target for structural studies because of health implications such as side effects of current antibiotics. While cryo-EM structures of human ribosomes have been recently reported to 4–5 Å resolution (18), no crystal structures are upcoming for now. Obtaining crystals of human ribosomes represents a major challenge, and any first clue on purification and crystallization conditions would be helpful, even if initial crystals usually diffract weakly. In a first strong effort in this direction, crystals of the human 80S ribosome were obtained using a nonstandard integrated structural biology approach, using a variety of established methods in a synergistic way rather than individually. This included establishing a detailed method for large-scale preparation of homogenous 80S ribosomes extracted from HeLa cells, which can be grown in large quantities and from which the ribosome purification is less complicated than from blood (18) or human placenta (29), majorly due to the ease of availability of HeLa cells and less tedious methods of initial lysis. Placenta handling requires immediate isolation of ribosomes, due to the inherent presence of large RNase contaminations; also, the lysis itself is tedious, requiring gauze filtration and dounce homogenizer owing to the presence of connective tissues. Similarly, blood samples require a ficoll-hypaque density-gradient centrifugation to separate out lymphocytes. Such extensive methods are not needed when handling HeLa cells: the usage of a detergent is sufficient for cell lysis without any additional purification step. Contaminants like ferritin in placenta are difficult to get rid of and tend to
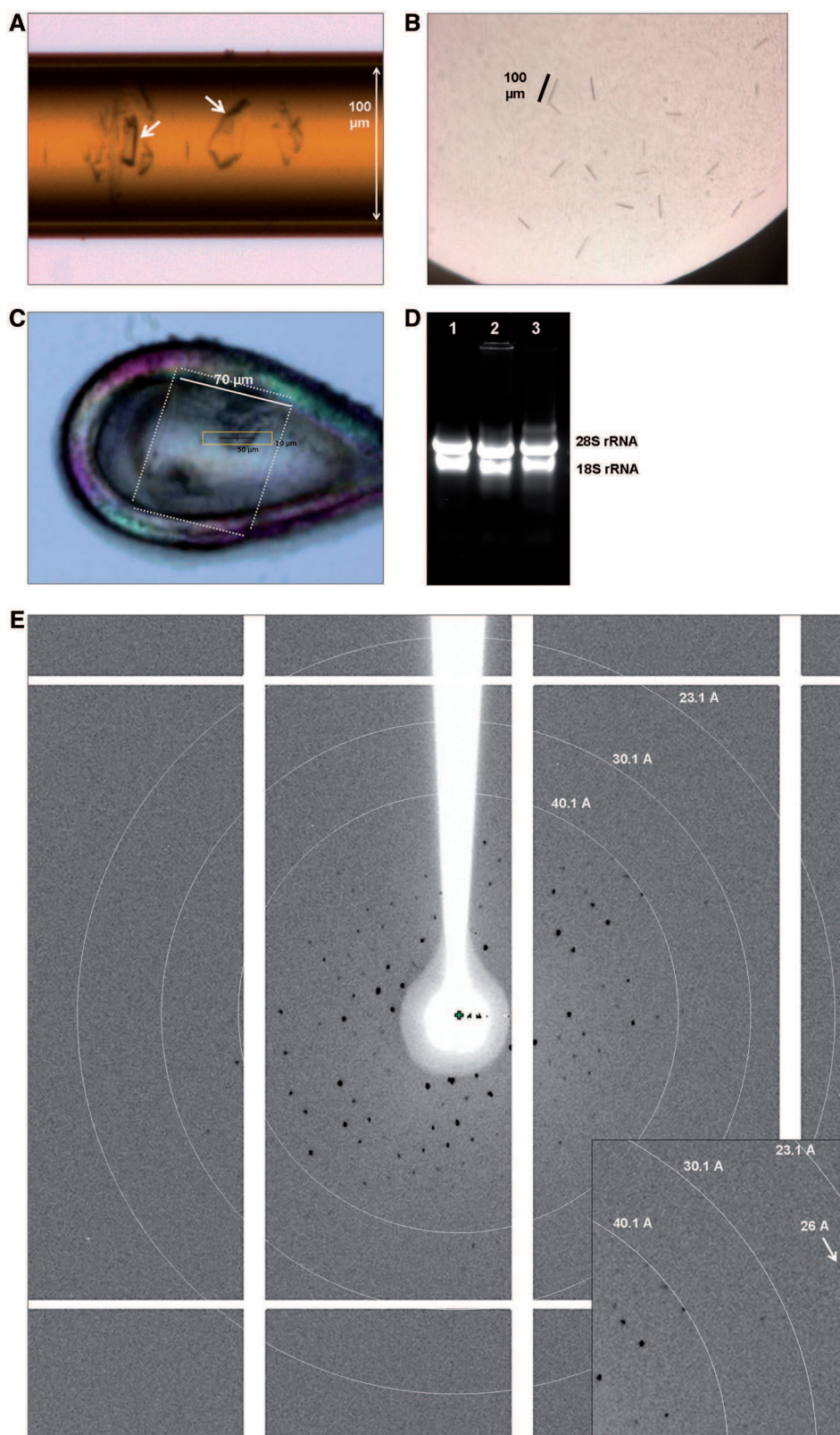
Figure 3. (**A**) First human 80S ribosome crystals obtained in the capillaries of a crystal harp. (**B**) Crystals with plate morphology reproduced in sitting drops that diffracted up to 26 Å, most are seen on the edge and thus give the impression of rods. (**C**) 80S crystal mounted in a cryogenic loop, tested for X-ray diffraction at the PX-II beamline at SLS. (**D**) Agarose gel depicting the sample stability under crystallization conditions as monitored by the presence of 18S and 28S rRNA. Lane 1 shows control ribosomes stored at −80°C, lanes 2 and 3 show ribosomes incubated with or without precipitant, respectively, for 4 weeks at 4°C. (**E**) The diffraction pattern shows a full reciprocal lattice with cell parameters of approximately a = 406 Å, b = 785 Å, c = 977 Å, with resolution rings indicated at 23, 30 and 40 Å. The inset shows diffraction spots extending to 26 Å resolution.

Table 2. Data collection statistics of a full data set collected from two regions of a single crystal at the SLS PX II beamline

X-ray data collection	Human 80S
Beamline	PX II (SLS)
Space group	C222(1)
Cell dimensions	
a, b, c (Å)	406.41, 784.99, 976.95
angles (°)	90, 90, 90
Resolution (Å)	150–42.0
Last shell	(43.9–42.0)
R_{sym} (%)	7.5 (74.1)
Reflections	5022
Completeness (%)	96.8 (88.6)
Redundancy	4.1 (3.2)
I/σ_i	7.7 (1.8)

Numbers in brackets refer to the highest resolution shell. R_{sym} is defined according to *XDS* (49).

be co-purified with ribosomes. Ribosomes from HeLa cells analyzed in the present study neither have these contaminants nor these RNA and protein degradation issues known for placental ribosomes. The next two steps following cell lysis in the purification procedure are usually used for standard ribosome purifications: sucrose cushion and a gradient to obtain homogenous 80S ribosomes, which leads to an additional sucrose removal step. Here, human 80S were purified using PEG20K precipitation, which is a more straightforward method [used also for yeast ribosomes (7)]. The detailed standardized purification protocol described in this work is a prerequisite for structural studies of the human ribosome and has allowed obtaining first crystals that are highly reproducible from various cell culture batches. A typical feature of the work presented here is the biophysical characterization by SEC-MALLS, AUC and includes the usage of sucrose gradients. For example, SEC-MALLS and AUC helped to check the sample solubility and stability in different ionic conditions, along with molecular weight detection to check for presence of aggregates or dissociated subunits. Also, the exploration of various purification procedures, optimized by monitoring sample purity and homogeneity (in particular absence of aggregation) through direct nondenaturing cryo-EM visualization of the ribosome particles in conjunction with crystallization trials is a prominent feature of this work and has been the key in obtaining first 80S crystals, thanks to the unprecedented combined usage of cryo-EM and crystallography. Rather than performing a detailed structural analysis, cryo-EM was used for examining the sample composition (presence of structured rRNA's and ribosomal proteins, presence of tRNA's) and structural integrity from quickly obtainable medium-resolution cryo-EM maps (using data sets 10 times smaller than used for high-resolution structures reported previously) with the aim of obtaining crystallizable material. This structural analysis revealed the presence of E-site tRNA in human 80S ribosomes depending on a subtle difference in the purification protocol, that is the salt concentration used during Puromycin treatment. The salt-sensitivity suggests that this E-site tRNA

binds in a nonspecific manner (50), probably also because no mRNA is present. This provides a valuable tool for obtaining either completely empty or E-site tRNA bound ribosomes, which can be used separately for further complex reconstitutions. For crystallization, we show that the presence of the E-site tRNA in human 80S ribosomes is helpful if not required because crystals could not be obtained with completely empty 80S ribosomes. This is probably because the E-site tRNA stabilizes the local ribosome structure, notably in the regions of the L1 and P-stalk proteins (51). The resulting conformation is compatible with the formation of inter-particle contacts in the crystal, as illustrated by the observed crystal nucleation and crystal growth, but different conformations may still be present in the asymmetric unit, as observed for yeast (7) and *T. thermophila* ribosomes (48). A related interesting aspect is that crystals can be obtained despite the presence of huge RNA ES elements, which protrude from the core structure (52); this is remarkable in terms of crystal packing, suggesting that ESs do not hamper crystallization in a significant way.

The use of capillaries for crystallization of large macromolecular complexes has been underexploited in the recent past, even though it is known to work well for smaller proteins (53,54). This work shows that unconventional methods like these can be helpful for initial screening of crystallization conditions, especially for inherently challenging samples and provide first hits to be further pursued in thicker capillaries or crystallization drops. The morphology of the human 80S ribosome crystals that form thin plates resembles that of early crystals of *T. thermophilus* ribosomes (55,56). After stabilization to reduce the fragile character of the thin crystals, they could be mounted in cryo loops in a rather straightforward manner. The crystals obtained diffract X-rays at a high-brilliance synchrotron source to up to 26 Å resolution, and show clear and fine diffraction spots along the reflection series of the diffraction pattern. The clean diffraction pattern also reveals that the crystals are monocrystalline, i.e. the crystal plates obtained under the crystallization conditions described here do not form multiple layers of distinct crystals, a phenomenon that otherwise occurs rather often for crystals with plate morphology. The diffraction pattern could be indexed, revealing that the crystals belong to the centric orthorhombic Bravais lattice type (space groups C222 or C2221, which have eight asymmetric units). The cell parameters of approximately $a = 406$ Å, $b = 785$ Å and $c = 977$ Å clearly indicate the presence of a large complex compatible with the Mw = 4.3 MDa estimated from the AUC and SEC-MALLS data. The calculation of the Matthews coefficient gives a value of $V_m = 3.0$ with three molecules in the asymmetric unit and a solvent content of 59%, or $V_m = 2.3$ with four molecules and 45% solvent [comparable with other ribosome crystals (57,58); the third alternative would give $V_m = 4.5$ with two molecules and a rather high solvent content of 73%]. The presence of three or four molecules in the asymmetric unit may be useful for noncrystallographic symmetry averaging, provided the ribosome conformations are sufficiently similar. The small clean

diffraction spots help in separating the spots and integrating the data, especially in the more challenging b- and c-directions. Because of the large cell parameters, fine-slicing was used with angular increments of 0.1–0.2° per frame in combination with a high-sensitivity and low readout noise pixel detector (Pilatus). In conclusion, the crystals are of excellent quality and diffraction data could be collected and processed, providing a first full data set at 42 Å resolution (Table 2). Their diffraction power, however, is limited (best resolution seen is 26 Å up to now, Figure 3E) most likely because of the little thickness of the crystals, which varies between 5 and 10 μm. In the future, it might be worth exploring conditions that promote crystal growth in the third dimension. Possibilities worth considering comprise further refined crystallization conditions such as the usage of additive screens and detergents, or removal of surface-exposed parts that may hamper further crystal growth such as loosely bound ribosomal proteins or part of the ES elements whose removal could help in obtaining larger crystals (some proteins could be removed by differential salt-wash as illustrated for the E-site tRNA, a procedure used for S1-removal; (59, 60) or cell line engineering such as done for L9-removal (61), which may be difficult to apply to human cells though. Although serum starvation was used (to synchronize cells), the ribosomes do not have any starvation factors bound, as compared with yeast ribosomes (7) or ribosomes isolated from blood (18), which have Stm-1 and Stm-1 like factors bound, respectively. This can be an advantage when forming translation complexes with mRNA and ribosomal translation factors for future crystallization or cryo-EM studies. Another aspect which we begun to analyze here is to refine the procedure for crystal stabilization and cryo-protection and reduce crystal mosaicity; controlled dehydration may also improve the diffraction of the crystals (7). In any case, while the crystals described here are diffracting relatively poorly and will need to be improved to reach a stronger diffraction power, they are the very first human ribosome crystals reported to our knowledge. This shows that the optimized purification procedure described herein provides crystallizable human ribosomes, a *sine-qua-non* condition for further crystallography work. The present work thus paves the way for future high-resolution crystal structures of the human ribosome, in isolated form or as complexes with mRNA, tRNAs and translation factors, with a major potential impact for studying molecular mechanisms and exploring medical applications in the ribosome field.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Meitian Wang and Florian Dworkowski from the Swiss Light Source for their excellent support on the PX I, II and III beamlines, Jean-Francois Ménétret for discussions on cryo-EM

REFERENCES

1. Dube,P., Bacher,G., Stark,H., Mueller,F., Zemlin,F., van Heel,M. and Brimacombe,R. (1998) Correlation of the expansion segments in mammalian rRNA with the fine structure of the 80 S ribosome; a cryoelectron microscopic reconstruction of the rabbit reticulocyte ribosome at 21 A resolution. *J. Mol. Biol.*, **279**, 403–421.
2. Chen,I.J., Wang,I.A., Tai,L.R. and Lin,A. (2008) The role of expansion segment of human ribosomal protein L35 in nuclear entry, translation activity, and endoplasmic reticulum docking. *Biochem. Cell Biol.*, **86**, 271–277.
3. Becker,T., Franckenberg,S., Wickles,S., Shoemaker,C.J., Anger,A.M., Armache,J.P., Sieber,H., Ungewickell,C., Berninghausen,O., Daberkow,I. *et al.* (2012) Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature*, **482**, 501–506.
4. Klaholz,B.P. (2011) Molecular recognition and catalysis in translation termination complexes. *Trends Biochem. Sci.*, **36**, 282–292.
5. Jackson,R.J., Hellen,C.U. and Pestova,T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 113–127.
6. Klaholz,B.P., Myasnikov,A.G. and van Heel,M. (2004) Visualization of release factor 3 on the ribosome during termination of protein synthesis. *Nature*, **427**, 862–865.
7. Ben-Shem,A., Garreau de Loubresse,N., Melnikov,S., Jenner,L., Yusupova,G. and Yusupov,M. (2011) The structure of the eukaryotic ribosome at 3.0 A resolution. *Science*, **334**, 1524–1529.
8. Klinge,S., Voigts-Hoffmann,F., Leibundgut,M., Arpagaus,S. and Ban,N. (2011) Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science*, **334**, 941–948.
9. Rabl,J., Leibundgut,M., Ataide,S.F., Haag,A. and Ban,N. (2011) Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science*, **331**, 730–736.

10. Verschoor,A., Srivastava,S., Grassucci,R. and Frank,J. (1996) Native 3D structure of eukaryotic 80s ribosome: morphological homology with *E. coli* 70S ribosome. *J. Cell Biol.*, **133**, 495–505.

11. Armache,J.P., Jarasch,A., Anger,A.M., Villa,E., Becker,T., Bhushan,S., Jossinet,F., Habeck,M., Dindar,G., Franckenberg,S. *et al.* (2010) Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. *Proc. Natl Acad. Sci. USA*, **107**, 19748–19753.

12. Spahn,C.M., Jan,E., Mulder,A., Grassucci,R.A., Sarnow,P. and Frank,J. (2004) Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell*, **118**, 465–475.

13. Spahn,C.M., Beckmann,R., Eswar,N., Penczek,P.A., Sali,A., Blobel,G. and Frank,J. (2001) Structure of the 80S ribosome from Saccharomyces cerevisiae–tRNA-ribosome and subunit-subunit interactions. *Cell*, **107**, 373–386.

14. Chandramouli,P., Topf,M., Menetret,J.F., Eswar,N., Cannone,J.J., Gutell,R.R., Sali,A. and Akey,C.W. (2008) Structure of the mammalian 80S ribosome at 8.7 A resolution. *Structure*, **16**, 535–548.

15. Taylor,D., Unbehaun,A., Li,W., Das,S., Lei,J., Liao,H.Y., Grassucci,R.A., Pestova,T.V. and Frank,J. (2012) Cryo-EM structure of the mammalian eukaryotic release factor eRF1-eRF3-associated termination complex. *Proc. Natl Acad. Sci. USA*, **109**, 18413–18418.

16. Sengupta,J., Nilsson,J., Gursky,R., Spahn,C.M., Nissen,P. and Frank,J. (2004) Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. *Nat. Struct. Mol. Biol.*, **11**, 957–962.

17. Halic,M., Becker,T., Frank,J., Spahn,C.M. and Beckmann,R. (2005) Localization and dynamic behavior of ribosomal protein L30e. *Nat. Struct. Mol. Biol.*, **12**, 467–468.

18. Anger,A.M., Armache,J.P., Berninghausen,O., Habeck,M., Subklewe,M., Wilson,D.N. and Beckmann,R. (2013) Structures of the human and Drosophila 80S ribosome. *Nature*, **497**, 80–85.

19. Becker,T., Bhushan,S., Jarasch,A., Armache,J.P., Funes,S., Jossinet,F., Gumbart,J., Mielke,T., Berninghausen,O., Schulten,K. *et al.* (2009) Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science*, **326**, 1369–1373.

20. Taylor,D.J., Devkota,B., Huang,A.D., Topf,M., Narayanan,E., Sali,A., Harvey,S.C. and Frank,J. (2009) Comprehensive molecular structure of the eukaryotic ribosome. *Structure*, **17**, 1591–1604.

21. Becker,T., Armache,J.P., Jarasch,A., Anger,A.M., Villa,E., Sieber,H., Motaal,B.A., Mielke,T., Berninghausen,O. and Beckmann,R. (2011) Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. *Nat. Struct. Mol. Biol.*, **18**, 715–720.

22. Hashem,Y., des Georges,A., Fu,J., Buss,S.N., Jossinet,F., Jobe,A., Zhang,Q., Liao,H.Y., Grassucci,R.A., Bajaj,C. *et al.* (2013) High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. *Nature*, **494**, 385–389.

23. Yonath,A. (2005) Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation. *Annu. Rev. Biochem.*, **74**, 649–679.

24. Franceschi,F. and Duffy,E.M. (2006) Structure-based drug design meets the ribosome. *Biochem. Pharmacol.*, **71**, 1016–1025.

25. Carter,A.P., Clemons,W.M., Brodersen,D.E., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, **407**, 340–348.

26. Brodersen,D.E., Clemons,W.M. Jr, Carter,A.P., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2000) The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell*, **103**, 1143–1154.

27. Wilson,D.N., Harms,J.M., Nierhaus,K.H., Schlunzen,F. and Fucini,P. (2005) Species-specific antibiotic-ribosome interactions: implications for drug development. *Biol.Chem.*, **386**, 1239–1252.

28. Kannan,K. and Mankin,A.S. (2011) Macrolide antibiotics in the ribosome exit tunnel: species-specific binding and action. *Ann. N. Y. Acad. Sci.*, **1241**, 33–47.

29. Matasova,N.B., Myltseva,S.V., Zenkova,M.A., Graifer,D.M., Vladimirov,S.N. and Karpova,G.G. (1991) Isolation of ribosomal subunits containing intact rRNA from human placenta: estimation of functional activity of 80S ribosomes. *Anal. Biochem.*, **198**, 219–223.

30. Jan,E. and Sarnow,P. (2002) Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J. Mol. Biol.*, **324**, 889–902.

31. Tyulkina,L.G. and Mankin,A.S. (1984) Inhibition of ribonuclease contamination in preparations of T4 RNA ligase, polynucleotide kinase, and bacterial alkaline phosphatase with bentonite. *Anal. Biochem.*, **138**, 285–290.

32. Jacoli,G.G., Ronald,W.P. and Lavkulich,L. (1973) Inhibition of ribonuclease activity by bentonite. *Can. J. Biochem.*, **51**, 1558–1565.

33. Belin,S., Hacot,S., Daudignon,L., Therizols,G., Pourpe,S., Mertani,H.C., Rosa-Calatrava,M. and Diaz,J.J. (2010) Purification of ribosomes from human cell lines. *Curr. Protoc. Cell Biol.*, **Chapter 3**, Unit 3.40.

34. Blobel,G. and Sabatini,D. (1971) Dissociation of mammalian polyribosomes into subunits by puromycin. *Proc. Natl Acad. of Sci. USA*, **68**, 390–394.

35. Lebowitz,J., Lewis,M.S. and Schuck,P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci.*, **11**, 2067–2079.

36. Schuck,P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.

37. Ludtke,S.J., Baldwin,P.R. and Chiu,W. (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.*, **128**, 82–97.

38. Tang,G., Peng,L., Baldwin,P.R., Mann,D.S., Jiang,W., Rees,I. and Ludtke,S.J. (2007) EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.*, **157**, 38–46.

39. Saxton,W.O. and Baumeister,W. (1982) The correlation averaging of a regularly arranged bacterial cell envelope protein. *J. Microsc.*, **127**, 127–138.

40. Rosenthal,P.B. and Henderson,R. (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.*, **333**, 721–745.

41. van Heel,M. and Schatz,M. (2005) Fourier shell correlation threshold criteria. *J. Struct. Biol.*, **151**, 250–262.

42. Kuwano,M., Endo,H. and Ikehera,Y. (1973) Differences in RNA formation and polyribosome metabolism in serum-starved normal and transformed cells. *Cancer Res.*, **33**, 2965–2971.

43. Grassucci,R.A., Taylor,D.J. and Frank,J. (2007) Preparation of macromolecular complexes for cryo-electron microscopy. *Nat. Protoc.*, **2**, 3239–3246.

44. Shenvi,C.L., Dong,K.C., Friedman,E.M., Hanson,J.A. and Cate,J.H. (2005) Accessibility of 18S rRNA in human 40S subunits and 80S ribosomes at physiological magnesium ion concentrations—implications for the study of ribosome dynamics. *RNA*, **11**, 1898–1908.

45. Moore,M.N. and Spremulli,L.L. (1985) Effects of cations and cosolvents on eukaryotic ribosomal subunit conformation. *Biochemistry*, **24**, 191–196.

46. Sperrazza,J.M., Russell,D.W. and Spremulli,L.L. (1980) Reversible dissociation of wheat germ ribosomal subunits: cation-dependent equilibria and thermodynamic parameters. *Biochemistry*, **19**, 1053–1058.

47. Beckmann,R., Spahn,C.M., Frank,J. and Blobel,G. (2001) The active 80S ribosome-Sec61 complex. *Cold Spring Harb. Symp. Quant. Biol.*, **66**, 543–554.

48. Weisser,M., Voigts-Hoffmann,F., Rabl,J., Leibundgut,M. and Ban,N. (2013) The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat. Struct. Mol. Biol.*, **20**, 1015–1017.

49. Kabsch,W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.

50. Nierhaus,K.H. (2006) Decoding errors and the involvement of the E-site. *Biochimie*, **88**, 1013–1019.

51. Fei,J., Kosuri,P., MacDougall,D.D. and Gonzalez,R.L. Jr (2008) Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation. *Mol. Cell*, **30**, 348–359.

52. Melnikov,S., Ben-Shem,A., Garreau de Loubresse,N., Jenner,L., Yusupova,G. and Yusupov,M. (2012) One core, two shells: bacterial and eukaryotic ribosomes. *Nat. Struct. Mol. Biol.*, **19**, 560–567.

53. Lorber,B., Sauter,C., Theobald-Dietrich,A., Moreno,A., Schellenberger,P., Robert,M.C., Capelle,B., Sanglier,S., Potier,N. and Giege,R. (2009) Crystal growth of proteins, nucleic acids, and viruses in gels. *Prog. Biophys. Mol. Biol.*, **101**, 13–25.

54. Giege,R. and Sauter,C. (2010) Biocrystallography: past, present, future. *HFSP J.*, **4**, 109–121.

55. Yonath,A., Glotz,C., Gewitz,H.S., Bartels,K.S., von Bohlen,K., Makowski,I. and Wittmann,H.G. (1988) Characterization of crystals of small ribosomal subunits. *J. Mol. Biol.*, **203**, 831–834.

56. Yusupov,M.M., Yusupova,G.Z., Baucom,A., Lieberman,K., Earnest,T.N., Cate,J.H. and Noller,H.F. (2001) Crystal structure of the ribosome at 5.5 A resolution. *Science*, **292**, 883–896.

57. Gao,Y.G., Selmer,M., Dunham,C.M., Weixlbaumer,A., Kelley,A.C. and Ramakrishnan,V. (2009) The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science*, **326**, 694–699.

58. Ben-Shem,A., Jenner,L., Yusupova,G. and Yusupov,M. (2010) Crystal structure of the eukaryotic ribosome. *Science*, **330**, 1203–1209.

59. Clemons,W.M. Jr, Brodersen,D.E., McCutcheon,J.P., May,J.L., Carter,A.P., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2001) Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: purification, crystallization and structure determination. *J. Mol. Biol.*, **310**, 827–843.

60. Duval,M., Korepanov,A., Fuchsbauer,O., Fechter,P., Haller,A., Fabbretti,A., Choulier,L., Micura,R., Klaholz,B.P., Romby,P. *et al.* (2013) Escherichia coli Ribosomal Protein S1 Unfolds Structured mRNAs Onto the Ribosome for Active Translation Initiation. *PLoS Biol.*, **11**, e1001731.

61. Selmer,M., Gao,Y.G., Weixlbaumer,A. and Ramakrishnan,V. (2012) Ribosome engineering to promote new crystal forms. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 578–583.

Supplementary Table 1. Large subunit proteins of the human ribosome and their score as detected by MS.

Large subunit proteins			
Protein name	Old nomenclature	Uniprot ID	Score
L1	L10A	P62906	566
L2	L8	P62917	1018
L3	L3	P39023	256
L4	L4	P36578	300
L5	L11	P63913	1184
L6	L9	P32969	1408
L6e	L6	Q02878	145
L8e	L7A	P62424	265
L11	L12	P30050	72
L13	L13A	P40429	45
L13e	L13	P26373	83
L14	L23	P62829	1833
L14e	L14	P50914	76
L15	L27A	P46776	71
L15e	L15	P61313	93
L16	L10	P27635	198
L18	L5	P46777	552
L18e	L18	Q07020	115
L19e	L19	P84098	113
L20e	L18A	Q02543	176
L21e	L21	P46778	54
L22	L17	P18621	534
L22e	L22	P35268	218
L23	L23A	P62750	610
L24	L26	P61254	68
L24e	L24	P83731	11
L27e	L27	P61353	37
L28e	L28	P46779	20
L29	L35	P42766	18
L29e	L29	P47914	13
L30	L7	P18124	176
L30e	L30	P62888	175
L31e	L31	P62889	34
L32e	L32	P62910	116
L33e	L35A	P18077	16
L34e	L34	P49207	3
L36e	L36	Q9Y3U8	7
L37e	L37	P61927	-
L38e	L38	P63173	26
L39e	L39	P62891	-
L40e	L40	P62987	15
L41e	L41	P62945	-
L43e	L37A	P61513	128
L44e	L36A	P83881	22
P1/P2	LP1	P05386	131
P2	LP2	P05387	349
P0	LP0	P05388	97

The proteins marked '-' were not detected in the sample, most likely due to their low molecular weight.

Supplementary Table 2. Small subunit proteins of the human ribosome and their score as detected by MS.

Small subunit proteins			
Protein name	Old nomenclature	Uniprot ID	Score
S1e	S3A	P61247	1328
S2	SA	P08865	1022
S3	S3	P23396	353
S4	S9	P46781	45
S4e	S4	P62701	248
S5	S2	P15880	771
S6e	S6	P62753	932
S7	S5	P46782	1065
S7e	S7	P62081	181
S8	S15A	P62244	132
S8e	S8	P62241	166
S9	S16	P62249	42
S10	S20	P60866	20
S10e	S10	P46783	695
S11	S14	P62263	253
S12	S23	P62266	813
S12e	S12	P25398	154
S13	S18	P62269	57
S14	S29	P62273	-
S15	S13	P62277	43
S17	S11	P62280	66
S17e	S17	P0CW22	280
S19	S15	P62841	508
S19e	S19	P39019	95
S21e	S21	P63220	205
S24e	S24	P62847	76
S25e	S25	P62851	41
S26e	S26	P62854	37
S27e	S27	P42677	32
S28e	S28	P62857	84
S30e	S30	P62861	6
S31e	S27A	P62979	77
RACK1	RACK1	P63244	79

The proteins marked '-' were not detected in the sample, most likely due to their low molecular weight.

Mass spectrometry analysis assigns a score for each protein detected by comparison of theoretical data to the unknown protein. Hence, the scores validate the presence of the relevant proteins, but the stoichiometry of the proteins could not be determined in greater detail because the amounts of protease-digested peptides detected by MS analysis vary significantly (possibly due to a variable protease activity for different ribosomal proteins). It can therefore not be excluded that some loosely-bound ribosomal proteins are present with variable stoichiometry in the 80S complex.

Supplementary Table 1. Large subunit proteins of the human ribosome and their score as detected by MS.

Large subunit proteins			
Protein name	Old nomenclature	Uniprot ID	Score
L1	L10A	P62906	566
L2	L8	P62917	1018
L3	L3	P39023	256
L4	L4	P36578	300
L5	L11	P63913	1184
L6	L9	P32969	1408
L6e	L6	Q02878	145
L8e	L7A	P62424	265
L11	L12	P30050	72
L13	L13A	P40429	45
L13e	L13	P26373	83
L14	L23	P62829	1833
L14e	L14	P50914	76
L15	L27A	P46776	71
L15e	L15	P61313	93
L16	L10	P27635	198
L18	L5	P46777	552
L18e	L18	Q07020	115
L19e	L19	P84098	113
L20e	L18A	Q02543	176
L21e	L21	P46778	54
L22	L17	P18621	534
L22e	L22	P35268	218
L23	L23A	P62750	610
L24	L26	P61254	68
L24e	L24	P83731	11
L27e	L27	P61353	37
L28e	L28	P46779	20
L29	L35	P42766	18
L29e	L29	P47914	13
L30	L7	P18124	176
L30e	L30	P62888	175
L31e	L31	P62889	34
L32e	L32	P62910	116
L33e	L35A	P18077	16
L34e	L34	P49207	3
L36e	L36	Q9Y3U8	7
L37e	L37	P61927	-
L38e	L38	P63173	26
L39e	L39	P62891	-
L40e	L40	P62987	15
L41e	L41	P62945	-
L43e	L37A	P61513	128
L44e	L36A	P83881	22
P1/P2	LP1	P05386	131
P2	LP2	P05387	349
P0	LP0	P05388	97

The proteins marked '-' were not detected in the sample, most likely due to their low molecular weight.

Supplementary Table 2. Small subunit proteins of the human ribosome and their score as detected by MS.

Small subunit proteins			
Protein name	Old nomenclature	Uniprot ID	Score
S1e	S3A	P61247	1328
S2	SA	P08865	1022
S3	S3	P23396	353
S4	S9	P46781	45
S4e	S4	P62701	248
S5	S2	P15880	771
S6e	S6	P62753	932
S7	S5	P46782	1065
S7e	S7	P62081	181
S8	S15A	P62244	132
S8e	S8	P62241	166
S9	S16	P62249	42
S10	S20	P60866	20
S10e	S10	P46783	695
S11	S14	P62263	253
S12	S23	P62266	813
S12e	S12	P25398	154
S13	S18	P62269	57
S14	S29	P62273	-
S15	S13	P62277	43
S17	S11	P62280	66
S17e	S17	P0CW22	280
S19	S15	P62841	508
S19e	S19	P39019	95
S21e	S21	P63220	205
S24e	S24	P62847	76
S25e	S25	P62851	41
S26e	S26	P62854	37
S27e	S27	P42677	32
S28e	S28	P62857	84
S30e	S30	P62861	6
S31e	S27A	P62979	77
RACK1	RACK1	P63244	79

The proteins marked '-' were not detected in the sample, most likely due to their low molecular weight.

Mass spectrometry analysis assigns a score for each protein detected by comparison of theoretical data to the unknown protein. Hence, the scores validate the presence of the relevant proteins, but the stoichiometry of the proteins could not be determined in greater detail because the amounts of protease-digested peptides detected by MS analysis vary significantly (possibly due to a variable protease activity for different ribosomal proteins). It can therefore not be excluded that some loosely-bound ribosomal proteins are present with variable stoichiometry in the 80S complex.

3.1.1 High-resolution cryo-EM structure of the human 80S

The technical and software advancements in the field of cryo-EM as detailed in section 3.4 allowed us to pursue further structural analysis of these purified ribosomes. Using data collected at the in-house Titan Krios electron microscope equipped with Falcon II CMOS camera and Cs corrector, we were able to reconstruct 80S structure without any bound factors at resolution close to 5 Å (Fig. 35).
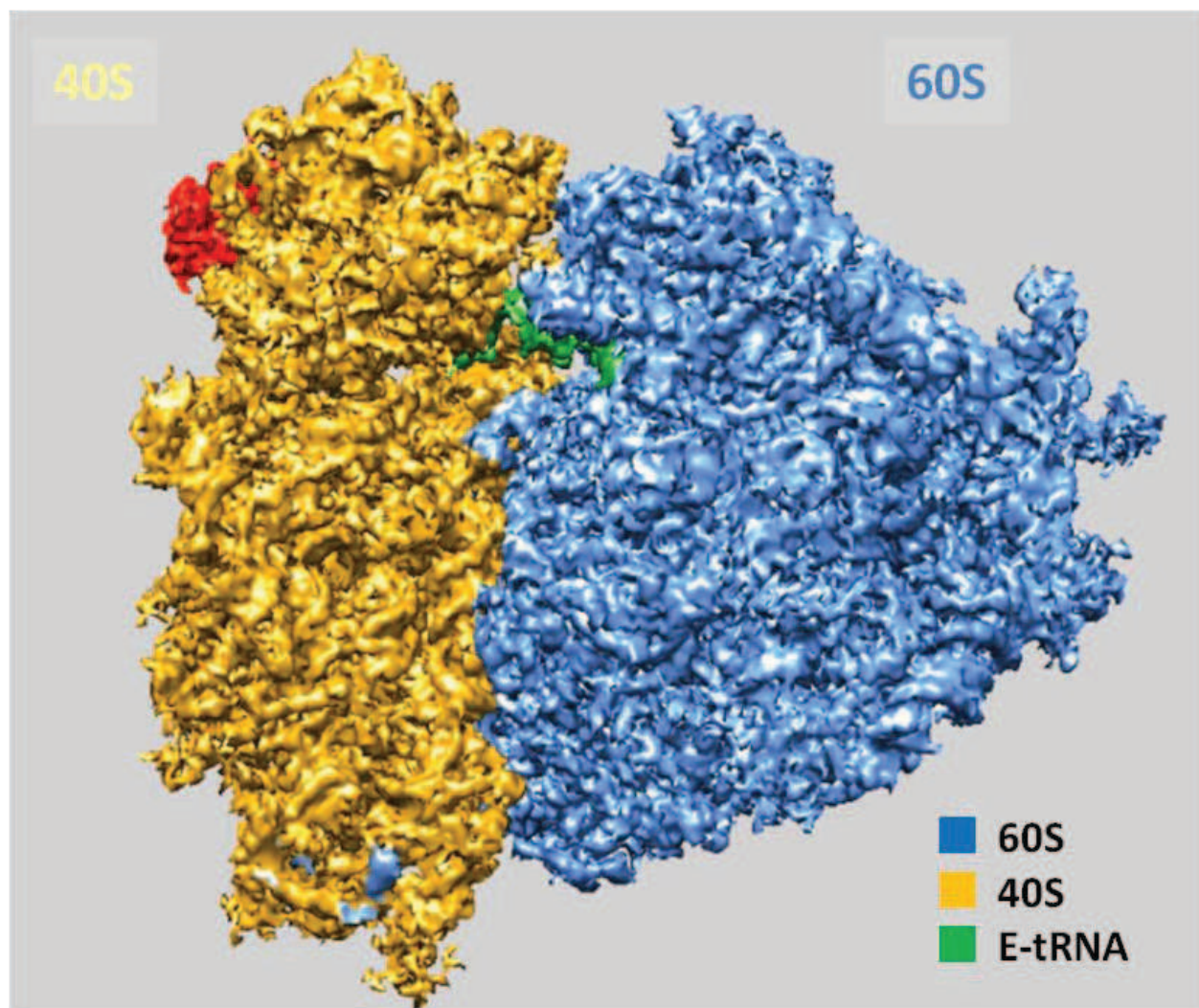
Figure 35: 5 Å resolution, reference structure of 80S.

At this resolution, we can observe most of the secondary structural elements of rRNA and proteins and even distinguish some bulky protein side chains like Phe and Trp. For example protein S6e, which is a single α-helical protein, can be resolved accurately as seen in the figure 35. Also, RACK1 complex, on the solvent-exposed side of the 40S, with six-fold pseudo symmetry can be easily recognised. Its seven-bladed β-propeller architecture has

each of its blades resolved at this resolution. The details observed in the density map are well-correlated with the resolution indicated by FSC (fourier shell correlation), obtained after Relion refinement procedure (Fig. 36). Presently, the structure is fit with model from 80S (Anger et al., 2013) using rigid-body fitting, further refinement and model building is ongoing.
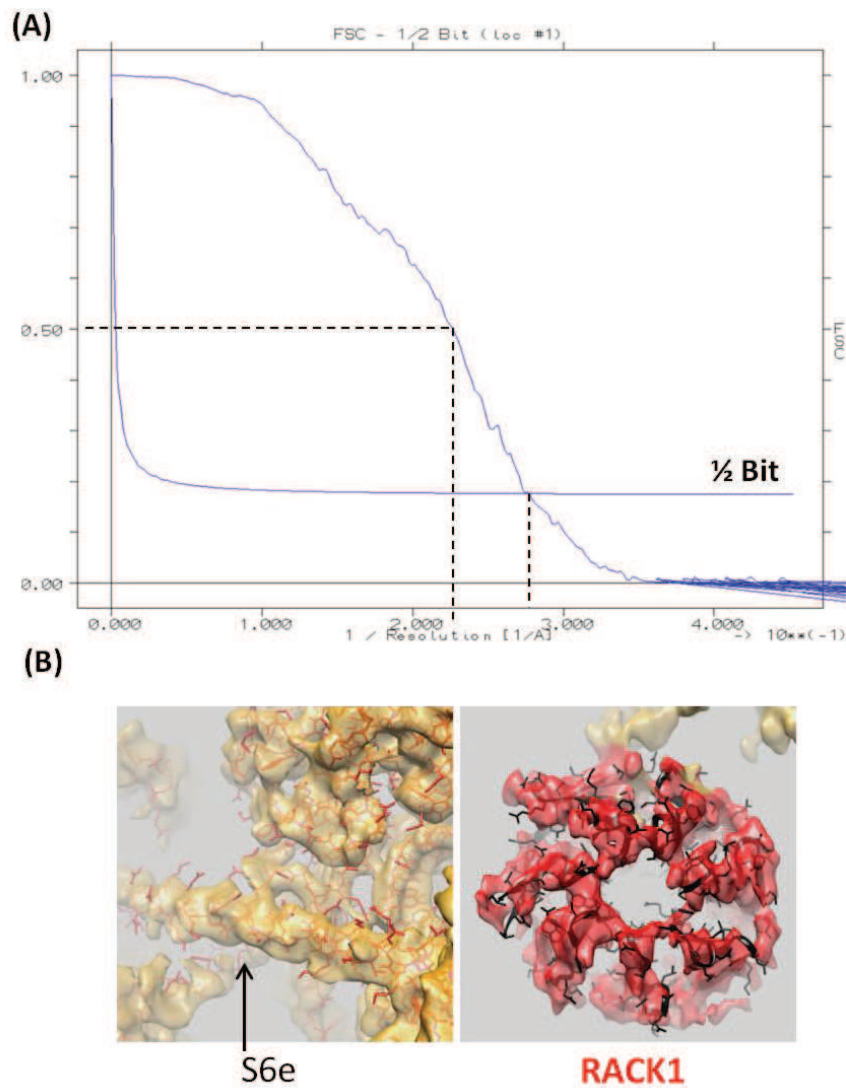
Figure 36: (A) The FSC curve showing the resolution attained according to 0.5 and ½ bit criteria. (B) The visibility of protein secondary structures is highlighted for protein S6e and RACK1.

Moreover, this structure could serve as a reference structure for future studies of protein translation stages, like in case of termination complexes, as explained below.

3.2 Termination factors

3.2.1 Cloning into bacterial expression vectors

eRF1 (class I release factor) and eRF3A (class II release factor) isoform I genes (NCBI access code NM_004730.3, NP_004721.1; NM_002094.3, NP_002085.2, respectively) were cloned in bacterial expression vectors. eRF1 was cloned using the Gateway system (LR reaction of pENTR-eRF1 and pCoGWA) into pCoGWA with an N-terminal six-residue histidine tag. The plasmid map and the site of insertion are marked in the Fig. 53 in Appendix.

eRF3 isoform I full-length has not yet been structurally characterized. Also, it has been shown that its expression is tedious in bacteria and must be solubilised by using thioredoxin (Kononenko, 2010). The gene corresponding to NP_002085.2 was chemically synthesised by GATC in pUC57 plasmid and double-digested with NdeI and BglII restriction enzymes. Alongside, empty pnEAvHX was cleaved with NdeI and BamHII. BglII and BamHI yield identical ends which allow the directional insertion of eRF3 (Fig. 37). The cleavage site in pnEAvHX was chosen such that thioredoxin and a six-residue long His tag lie at the N-terminal to allow expression and affinity purification.

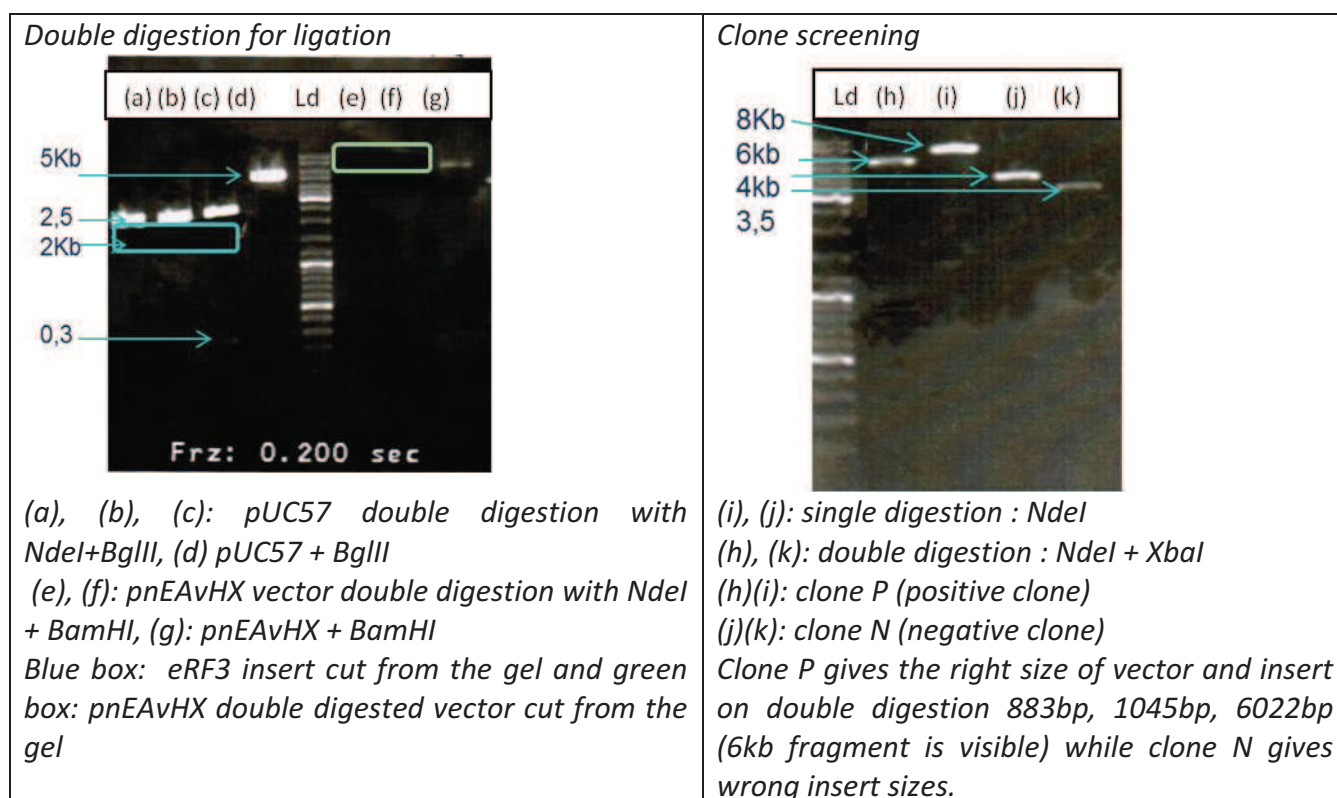| Double digestion for ligation	Clone screening
(a), (b), (c): pUC57 double digestion with NdeI+BglII, (d) pUC57 + BglII (e), (f): pnEAvHX vector double digestion with NdeI + BamHI, (g): pnEAvHX + BamHI Blue box: eRF3 insert cut from the gel and green box: pnEAvHX double digested vector cut from the gel	(i), (j): single digestion : NdeI (h), (k): double digestion : NdeI + XbaI (h)(i): clone P (positive clone) (j)(k): clone N (negative clone) Clone P gives the right size of vector and insert on double digestion 883bp, 1045bp, 6022bp (6kb fragment is visible) while clone N gives wrong insert sizes.

Figure 37: Agarose gels for eRF3 cloning in pnEAvHX vector.

Prior to the protein expression tests, the sequence of both clones was verified using sequencing at GATC Biotech Company. For both proteins, expression is under control of the T7 promoter which is in turn controlled by the lactose operon, induced by IPTG.

3.2.2 Protein expression and purification

A previously reported protocol was followed to purify eRF1 (Andjelkovic et al., 1996; Song et al., 2000). Plasmid was transformed into BL21 (DE3) cells and expression tests were performed for these cells at 37°C and 18°C. Cells grown in 2X LB media at 37°C, and induced overnight with IPTG at 18°C, gave the best yields of eRF1. The eRF1 overexpressed cells were lysed using sonication with the lysis buffer, and processed as explained below.
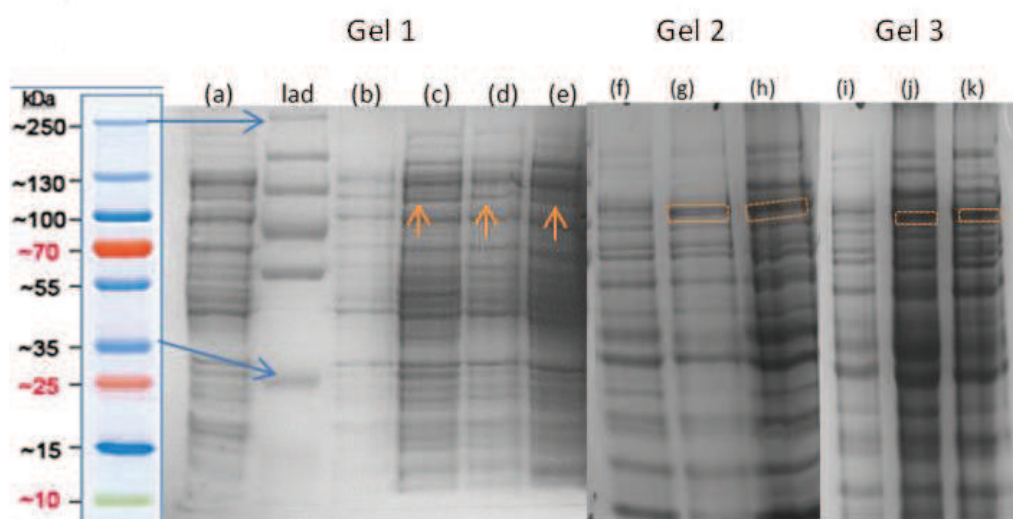
Figure 38: eRF3 expression tests on 10 % SDS PAGE gels. Gel 1 shows expression in BL21 (DE3) at different temperatures in 2X LB media. lad denotes the protein molecular weight ladder, (a) Culture before induction with IPTG, (b) culture after 4 hrs at 25°C, (c) after 16 hrs at 25°C (d) after 4 hrs at 37°C (e) after 16 hrs at 37°C. Gel 2 shows BL21 (RIL) transformed bacteria after 16 hrs at 37°C in different media. (f) Before induction, (g) auto-inducible media (h) 2X LB. Gel 3 BL21 (DE3) transformed bacteria after 16 hrs at 37°C in different media. (i) Before induction, (j) 2X LB (k) auto-inducible media.

Standardisation of the eRF3 purification (Kononenko et al., 2010) was more rigorous as it involved choosing the suitable competent cells for transformation, the temperature and medium for protein expression. Initially, the plasmid was transformed into BL21 (RIL) and BL21 (DE3) cells and the protein expression was tested at 25°C and 37°C after 4 hrs and 16 hrs of cell growth at 37°C for 16 hrs in two different culture media; 2X LB and auto-inducible medium (Fig. 38). Interestingly, the growth in auto-inducible medium did not really increase the production of eRF3, but the amount of contaminants present in the

soluble fraction was significantly decreased. Cell cultures at 37°C for 16 hrs provide a higher amount of cells expressing the protein at the size of about 100 kDa. Also, there is no formation of inclusion bodies or any degradation of the protein.

Usually, the cells were grown in batches of 6L per purification, but in order to obtain a higher yield both the proteins were expressed in fermentor of 20L. The fermentor (Infors HT Techfors) allows a better control over growth parameters, agitation, oxygenation and temperature. Since the protein purified is from the same cell culture batch, crystal reproducibility is also increased.
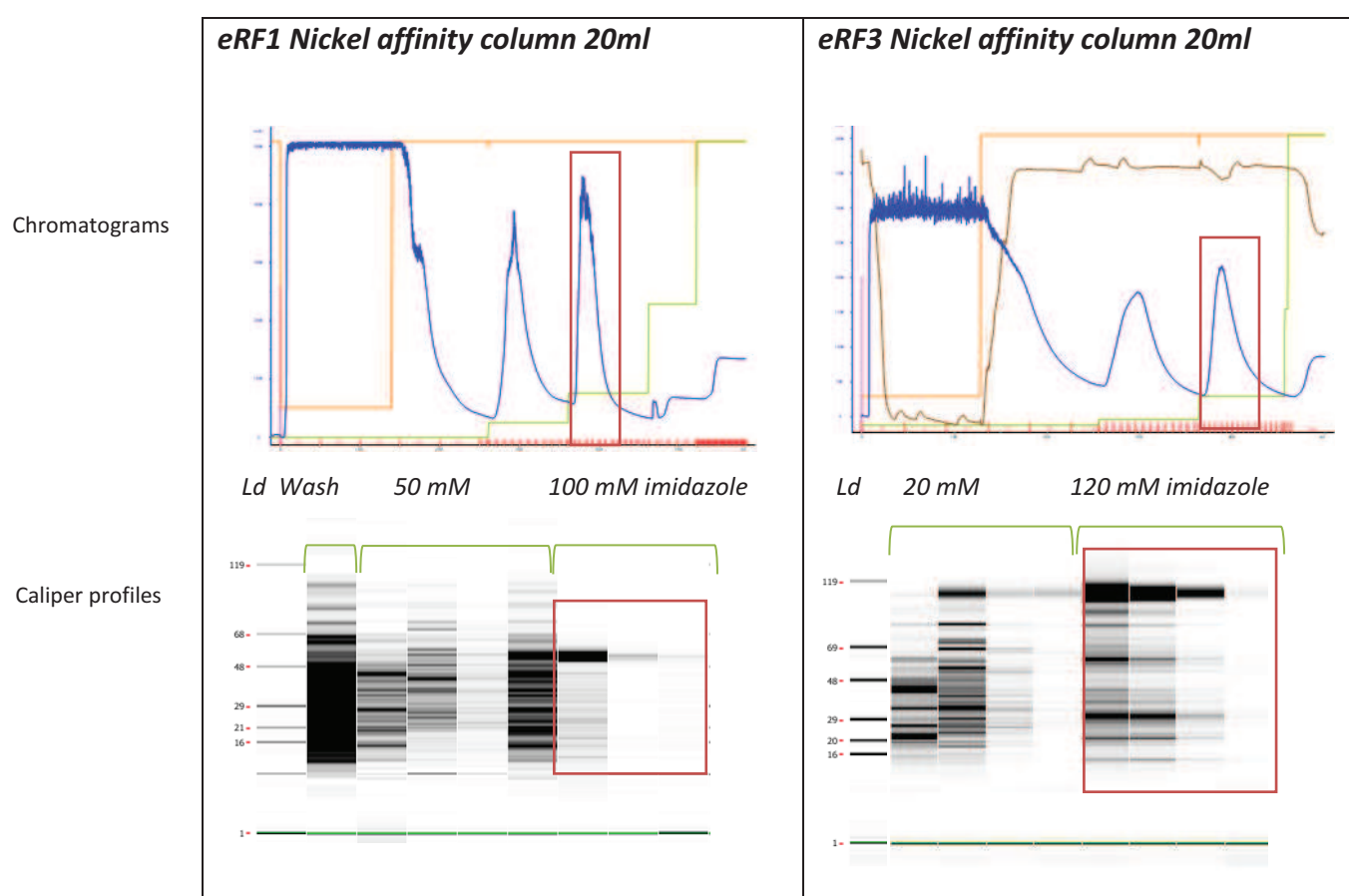
Figure 39: Chromatogram and gel profiles for affinity purification of eRF1 and eRF3. The gel profiles were obtained on LabChip GX Caliper®.

The cells were lysed using sonication, and especially since eRF3 is highly sensitive to degradation, necessary precautionary measures were taken to prevent overheating the sample. Both eRF1 and eRF3 contain a hexa-histidine residue tag at the N-terminus, so the protein was pulled out using Ni-NTA column. The protein was thereafter eluted with a gradient of imidazole (Fig. 39). In the first step of imidazole wash, possible contaminants

were removed (50 mM for eRF1 and 20 mM for eRF3). The second step of the gradient at 150mM imidazole eluted the protein. Fractions were collected and verified with SDS page analysis. The protein fractions were pooled and dialysed overnight, ideally against a volume 100 times the sample, to reduce salt to a minimum, in order to perform the next step of ion exchange (Fig. 40).
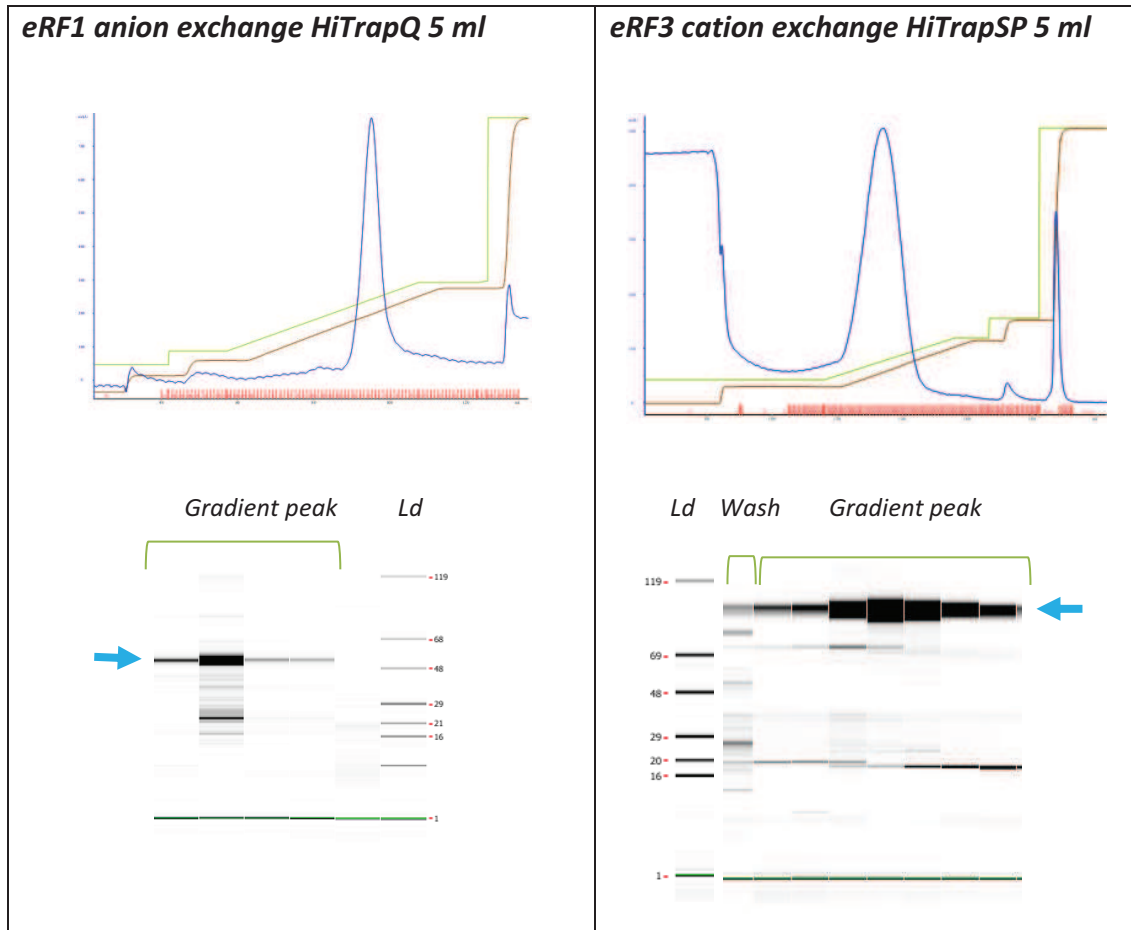
Figure 40: Chromatogram and gel profiles for ion exchange step for eRF1 and eRF3 purification.

eRF1 was purified with the anion exchange, however for eRF3 it was observed that cation exchange worked better in spite of a net negative charge on protein at pH 7.5. This could be because of lower K_d (dissociation constant) of full-length eRF3 for GDP association, as compared to eRF3 lacking the N-terminus. Also, the eRF1-eRF3-GTP complex has been observed to have considerably higher solvent accessible surface area as compared to the complex with eRF3 lacking the N terminus due to conformational changes (Kononenko et al., 2010). The same is true for the isolated eRF3, which might be responsible for the surface charge being different as eRF3 has an inherent dynamic structure. Both proteins were eluted between 150 mM and 200 mM salt concentration. For eRF1 purification, an extra

step of S75 gel filtration was performed to remove contaminants that can hinder the complex formation (Fig. 41). eRF1 requires the presence of 10% glycerol all throughout the purification for solubilisation, while eRF3 is stable and solubilised in buffer probably because of thioredoxin which probably helps in its folding and solubilisation. The yield of both factors is approximately 12-14 mg from 6L cultures. Once both proteins were purified to homogeneity, they were either frozen in aliquots or used directly to form the complex.

3.2.3 eRF1-eRF3 complex formation and biophysical characterization

Complex formation and tag removal was tried in two different methods. In the first method, the two proteins were incubated separately with TEV to remove the tag, followed by complex formation. The other method being where, eRF1 and eRF3 were incubated together with TEV to allow cleavage and complex formation.
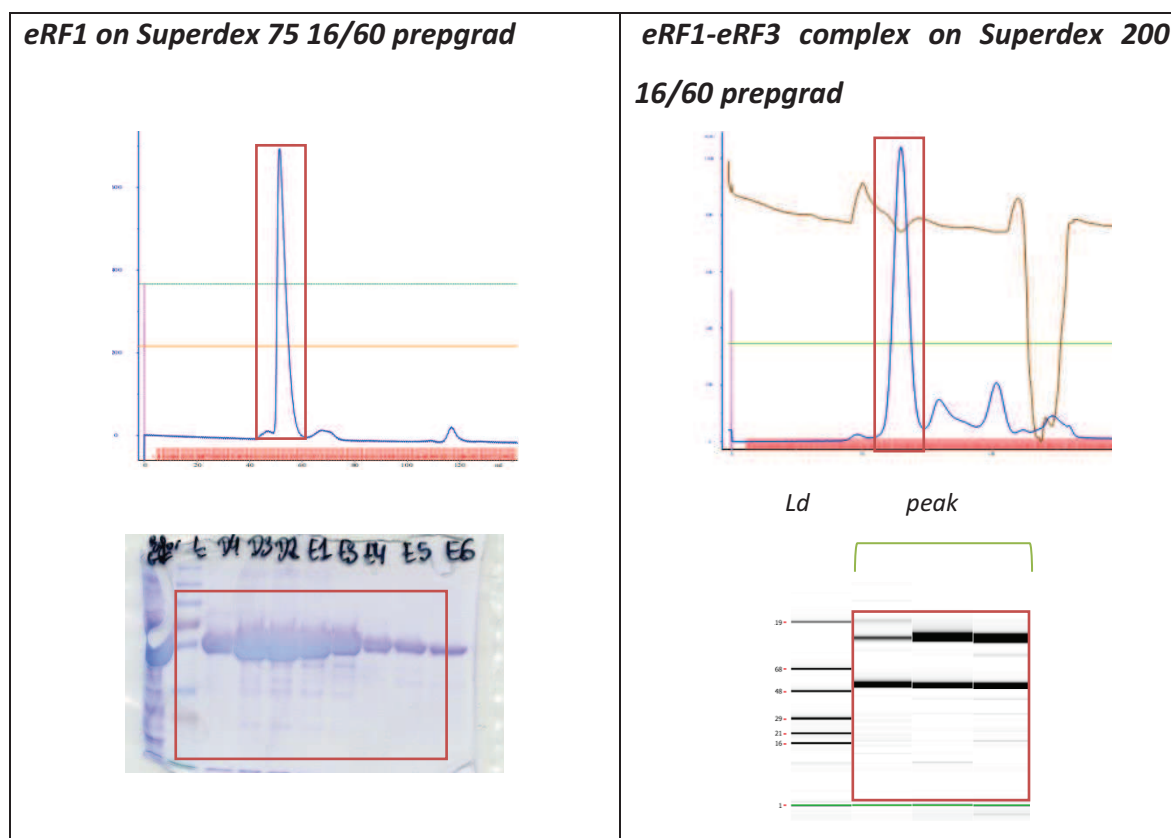
Figure 41: Chromatogram and 10% SDS gel for size exclusion chromatography of eRF1 (left panel) and complex of eRF1 and eRF3 (right panel). The red boxes indicate the peak fractions each corresponding to the purified protein.

The first method did not seem to be efficient in cleaving the tag, when purified eRF1 and eRF3 were incubated separately, with TEV for overnight incubation and followed by Ni-NTA column to remove the uncleaved protein. Less than 60% of protein was recovered

without the tag. But when eRF1 and eRF3 were incubated together with TEV overnight, almost 90% of total protein was recovered as a complex free of N-terminal tags. S200 gel filtration was important to separate the unbound protein and the cleaved tags from the protein complex (Fig. 41).

Mass spectrometry

ESI-TOF analysis was done under native and denaturing conditions to check for the protein complex, and ensure that it is stable in the buffer used. In native conditions, a peak at 118 kDa corresponds to the complex although there are peaks at the molecular weight of free proteins (50 kDa and 68 kDa) too (Fig. 42). Under denaturing conditions, both proteins were identified confirming the presence of the desired complex.
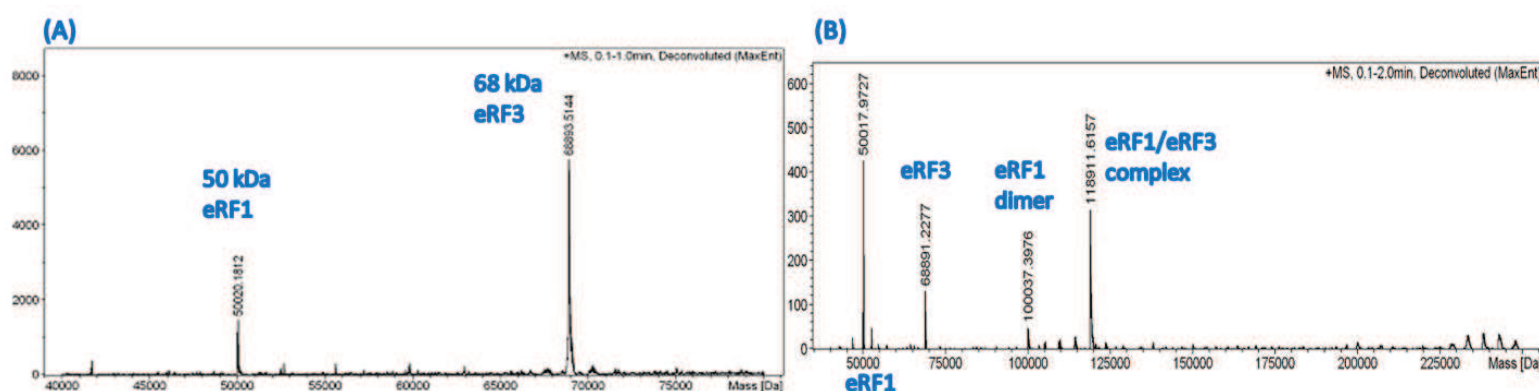
Figure 42: Mass spectrometry analysis of the complex (A) in denaturing conditions, (B) in native conditions.

Dynamic light scattering (DLS)

DLS estimates the hydrodynamic radius and the polydispersity of the sample and essentially revealing the homogeneity of the sample. The eRF1-eRF3 complex initially contains 10 % glycerol in the buffer, and glycerol is known to solubilise the sample instead of aiding crystallisation. Therefore, the complex was prepared in different buffers going from 10% to 0% glycerol, during S200 elution of the complex. Although no visible aggregates were observed even in the absence of glycerol, but the polydispersity of the sample varied between 23 and 7 (Fig. 43 and Table 8). Crystallisation plates were set up with 10%, 5% and 2.5% glycerol; crystals were obtained only in the sample containing 10% glycerol. This can be explained by the polydispersity curves obtained, since higher polydispersity indicates a heterogeneous sample with a large number of species, which could in turn prevent

116

crystallization. It should also be noted that the hydrodynamic radius increases on increasing glycerol concentration, as it interacts with hydrophobic surfaces of protein that favours amphiphilic interface formation and thus, prevents aggregation.
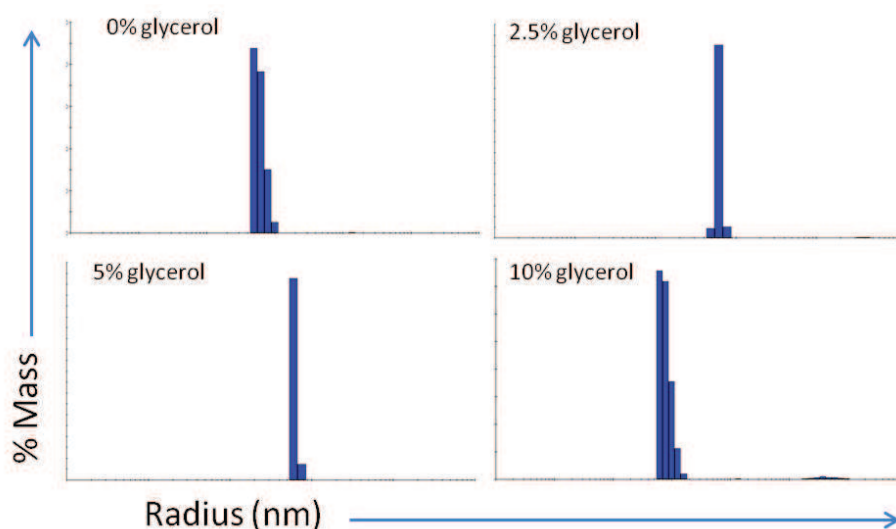
Figure 43: DLS profiles for eRF1-eRF3 protein complex.

Table 8: DLS analysis of the complex, denoting the polydispersity with decreasing glycerol content in the buffer.

% glycérol	Rh (nm)	% Polydispersity (Pd)	% Mass
10%	9.4	23.0	99.8
5%	6.138	8.5	99.7
2.5%	6.241	7.4	99.5
0%	6.038	20.4	99.5

Native gel analysis

Native gel analysis readily displays the complex formation and protein stability as a shift of molecular weight with respect to the controls. In the absence of denaturing agents like SDS, the proteins migrate on the poly acrylamide gel according to their net charge and conformation. eRF1 is known to exist as a dimer (100 kDa) in solution (Cheng et al., 2009), and as seen on the 6% gel, eRF3 (68 kDa) migrates faster than eRF1. The eRF1-eRF3 complex has a band similar to the purified eRF1 because the molecular weight for complex is 118 kDa and for an eRF1 dimer it would be around 100 kDa. But the complex formation is ensured in 1:1 molar ratio, since there is no free eRF3 visible in this sample (Fig. 44). Same holds true for the eRF1-eRF3 complex with the GTP-analogue, GMPPCP.

117
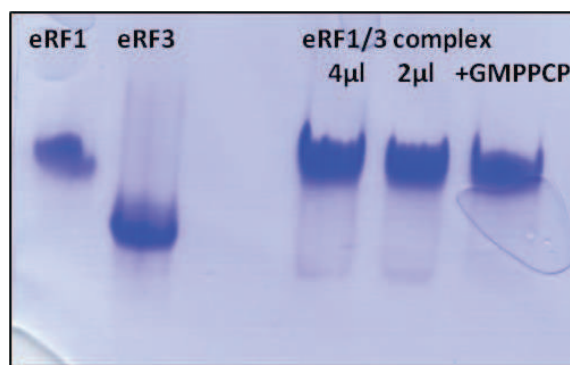
Figure 44: Native gel depicting the eRF1-eRF3 complex (118 kDa) in comparison with only eRF1 (100 kDa dimer in solution), and only eRF3 (68 kDa).

Microscale Thermophoresis Technology (MST)

To check for interactions of the two release factors in solution and determine the K_d, as has been reported in literature for eRF1-eRF3 complex (Kononenko et al., 2010), MST (Nanotemper) was used. Isothermal titration calorimetry (ITC) has been the technique used for such studies earlier, but MST offers a major advantage of using at least ten times less sample. However, the standardisation process for the suitable buffers to get a good signal to noise ratio, is equally time consuming for both.
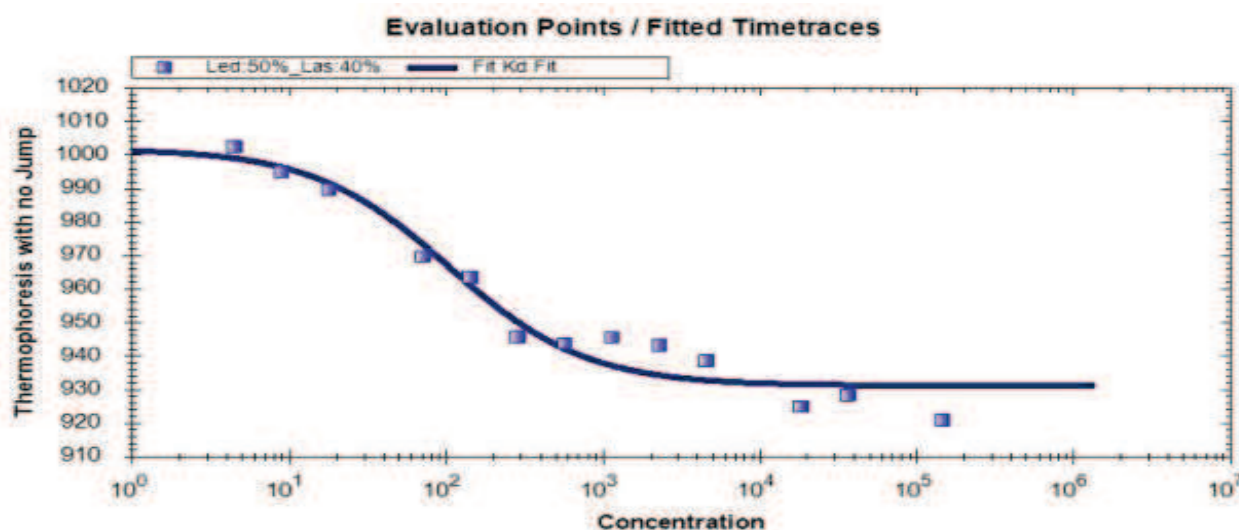
Figure 45: MST analysis of eRF1 and eRF3 complex to determine K_d.

eRF1 was serial-diluted in S200 gel filtration buffer, into 16 tubes, starting with the highest concentration. 100 µl of 20 µM eRF3 was labelled at either lysine or cysteine residues with FITC flour and diluted to 1:20 or 1:50 ratio in water and checked for signal, if

118

the signal was in the right range to be detected by the Nanotemper device. A final concentration of 32 nM was mixed with the serial-diluted eRF1 series and centrifuged to remove any aggregates before loading into capillaries for Nanotemper measurements.

A negative control was measured first, with only the fluorescent eRF3 to verify the non-specific interactions with capillary or buffer. eRF3 in absence of 0.05% Tween-20 failed to show any constant fluorescence (a variation of up to 20 units is acceptable), which is an indication of non-specific interactions occurring in the capillary. So to avoid these non-specific interactions, 0.05% Tween-20 was added to buffer for eRF1 as well as for eRF3. The complex fluorescence measurement gave a K_d of 150 nM (Fig. 45) which is in the range observed in previous studies, 200 nM (Kononenko et al., 2010).

3.2.4 Crystallization of the eRF1-eRF3 complex

After confirming the eRF1-eRF3 complex homogeneity, I proceeded with complex crystallization. The initial tests were performed with conditions that yielded crystals of eRF1 (Song et al., 2000), eRF3 (Kong et al., 2004) and the truncated eRF1-eRF3 complex missing the N-terminus and G domain of eRF3 (Cheng et al., 2009). Not surprisingly, these conditions failed to provide crystals, probably because the presence of the N-terminus and G domain changes the net surface charge of the complex that possibly altered the precipitation rate for the protein complex. So, all the 16 commercial screens available at the crystallization platform at IGBMC were tested with different sample preparations, glycerol contents and concentrations ranging from 9-16 mg/ml.

The initial hits obtained were tested at the in-house X-ray source and SLS synchrotron that allowed us to narrow down to the conditions mentioned in Table 9 for pursuing further optimisation of crystallization conditions. Crystals obtained using condition B diffracted to 20 Å which was good enough to derive the space group and cell parameters. The crystal belongs to C2 space group with a=135 Å, b=75 Å, c=80 Å, $\alpha=\gamma=90°$ and $\beta=118°$. Hanging and sitting drop setups were used to reproduce crystals in those conditions with larger drop volumes (1.5+1.5 µl). Simultaneously, additives screen from Hampton Research were used in 2% and 1% with the original condition to optimize and obtain good diffraction quality crystals.

Table 9: The crystallisation conditions for initial hits of eRF1-eRF3 protein complex with 10% glycerol in GF buffer and various precipitant agents.
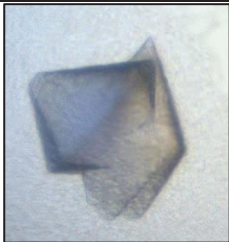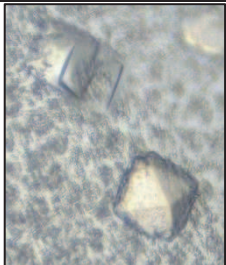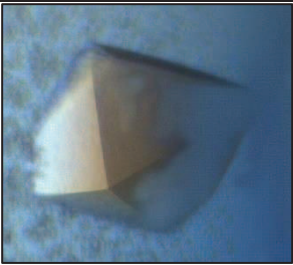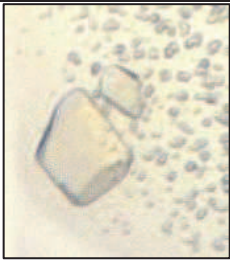
Condition	Condition A	Condition B	Condition C
Buffer	50 mM Sodium cacodylate, pH 6.5	50 mM Sodium cacodylate, pH 7	50 mM Sodium cacodylate, pH 6
Precipitant	2M $(NH_4)_2SO_4$	10% 2-Methyl-2-propanol	5% PEG 4000
Salt	10 mM $MgSO_4$	10 mM $MgCl_2$	20 mM $MgCl_2$
Additives		1 mM Spermidine	2.5 mM Spermine Tetrahydrochloride

For freezing the crystals, different cryoprotectants were screened including paraffin oil, silicon oil, ultimate cryoprotectant (UCP- 8% Glycerol, 8% ethylene glycol, 9% sucrose, 2% glucose) and 20-30% glycerol. Essentially, the best results were obtained with UCP for condition B and C. Condition B was particularly tricky to handle because of the volatile *tert*-butanol. The organic solvent evaporates on opening the drop, leaving the crystal in a state prone to damage. The initial data sets for the crystals obtained using condition C were collected to a resolution of around 10 Å to get a preliminary map to verify the presence of both the proteins. However, it was found by molecular replacement using the individual eRF1 and eRF3 crystal structures (PDB codes 1DT9 and 3J5Y) that only one protein was present in this condition and the resolution did not improve upon varying the cryoprotectants or adding additives in the crystallization condition. Besides, condition A has 2M ammonium sulphate which precipitates easily, making the crystals difficult to freeze and can also possibly dissociate the complex because of the high salt concentration. So these conditions were not pursued further.

Instead, the complex without glycerol in GF buffer was followed. Crystals were obtained in conditions listed in Table 10. Biochemically, the role of glycerol can be interpreted as follows. Glycerol interacts with large hydrophobic patches that form an amphiphilic interface between the hydrophobic surface and the polar solvent. This prevents protein aggregation/precipitation and instead solubilises the protein (Vagenende et al., 2009). Thus, depending upon the protein complex being investigated glycerol can inhibit

crystallization, since it would rather keep the protein in a soluble state. The same behaviour possibly occurs with the eRF1-eRF3 complex as well.

Table 10: The crystallisation conditions for initial hits of eRF1-eRF3 protein complex in the absence of glycerol in the GF buffer.

Condition	Condition D	Condition E	Condition F	Condition G
Buffer	50 mM Sodium cacodylate, pH 6.0	50 mM Sodium cacodylate, pH 6.5	50 mM Sodium cacodylate, pH 6.5	50 mM Sodium cacodylate, pH 6.0
Precipitant	10% PEG 400	1.3M Li_2SO_4	1.3M Li_2SO_4	1.7M $(NH_4)_2SO_4$
Salt	15 mM $MgCl_2$	10 mM $Mg(OAc)_2$	30 mM $MgCl_2$	15 mM $Mg(OAc)_2$
Additives	3 mM Spermine Tetrahydrochloride		1 mM Spermine Tetrahydrochloride	
Crystal				

The best diffracting crystals were obtained with 1.3 M Li_2SO_4 as a precipitant in condition F. The crystals were frozen using a higher concentration of Li_2SO_4 (1.8M) as a cryoprotectant. There was a thick skin on the drop, probably caused by precipitation, which was removed before freezing the crystal. Full data set of this crystal, diffracting to 3.7 Å resolution was collected (Table 11) and processed using molecular replacement strategy. The crystal belongs to $P4_12_12$ space group with a=105.7 Å, b=105.7 Å, c=218.49 Å, $\alpha=\beta=\gamma=$ 90°, and possesses 44 % solvent content with Matthews coefficient of 2.2.

Structure solution was obtained with molecular replacement using eRF1 (PDB code 1DT9) and eRF3 (PDB code 3E1Y and 3J5Y) as search models. The Matthews coefficient suggested that the asymmetric unit contains two molecules of eRF1-eRF3 complex, or slightly less such as two copies of eRF1 and a single copy of eRF3. Initially, MR searches were performed systematically with varying the number of copies and combinations of eRF1 and eRF3 structures. Finally, a rational approach with independent domain searches was

performed, that resulted in structural solution. First the N and C domains of eRF1 were searched and fixed in position, followed by the search for flexible M domain which resulted in placement of the first molecule of eRF1 (referred to as form I – green in Fig. 46). Thereafter, the N domain of second eRF1 was placed successfully. The M and C domains of the second eRF1 were searched as a single chain. This second molecule of eRF1 is referred to as form II (blue in Fig. 46). Once both the eRF1 molecules were localised, then β-barrels domain II and III were searched. Both these domains being rich in β-sheets have weak electron density. The G domain was placed the last, resulting in the final solution with two copies of eRF1 and one copy of eRF3 in an asymmetric unit. The LLG, TFZ score were reported to be 1306 and 37.5 respectively by Phaser (CCP4 suite).

Table 11: Data collection and refinement statistics. Values in parentheses indicate the values in the outer shell.

Data collection	Human eRF1-eRF3 complex
Wavelength (Å)	0.9999
Resolution limit (Å)	3.9-105
Space group	$P4_12_12$
Cell parameters	a=105.7, b=105.7, c=218.15, $\alpha=\beta=\gamma=90°$
Total reflections	43,872
Unique reflections	10,739
I/σ	15.1 (1.2)
Completeness (%)	93.2 (95.4)
R_{merge} %	0.038 (0.847)
Multiplicity	4.1 (4.3)
Refinement statistics	
Data range (Å)	42-4
Used reflections	11,456
R_{work} (%)	44
R_{free} (%)	46
RMSD	
Bond length	0.008
Bond angle	0.96

The Y-shaped eRF1 structure with the N, C and M domains is well resolved, but the loops present in the minidomain and the C domain are not. In the case of eRF3, the G domain and the two β-barrel motifs, domain II and III are well resolved. But in the molecular replacement search, the position and orientation of the 213 residues long eRF3 N domain could not be found. It could be because this domain tends to be unstructured in nature. The

G domain was observed to interact with the N terminal domain of form I eRF1 molecule in the asymmetric unit (Fig. 46) and domain II interacts with the M-domain of the form II eRF1 molecule. Also, there is no nucleotide bound in this complex, as the GTP binding pocket, residues 217-231, does not possess any extra positive density. Moreover, the switch helix residues of the eRF3 G domain were observed in an unstructured form.
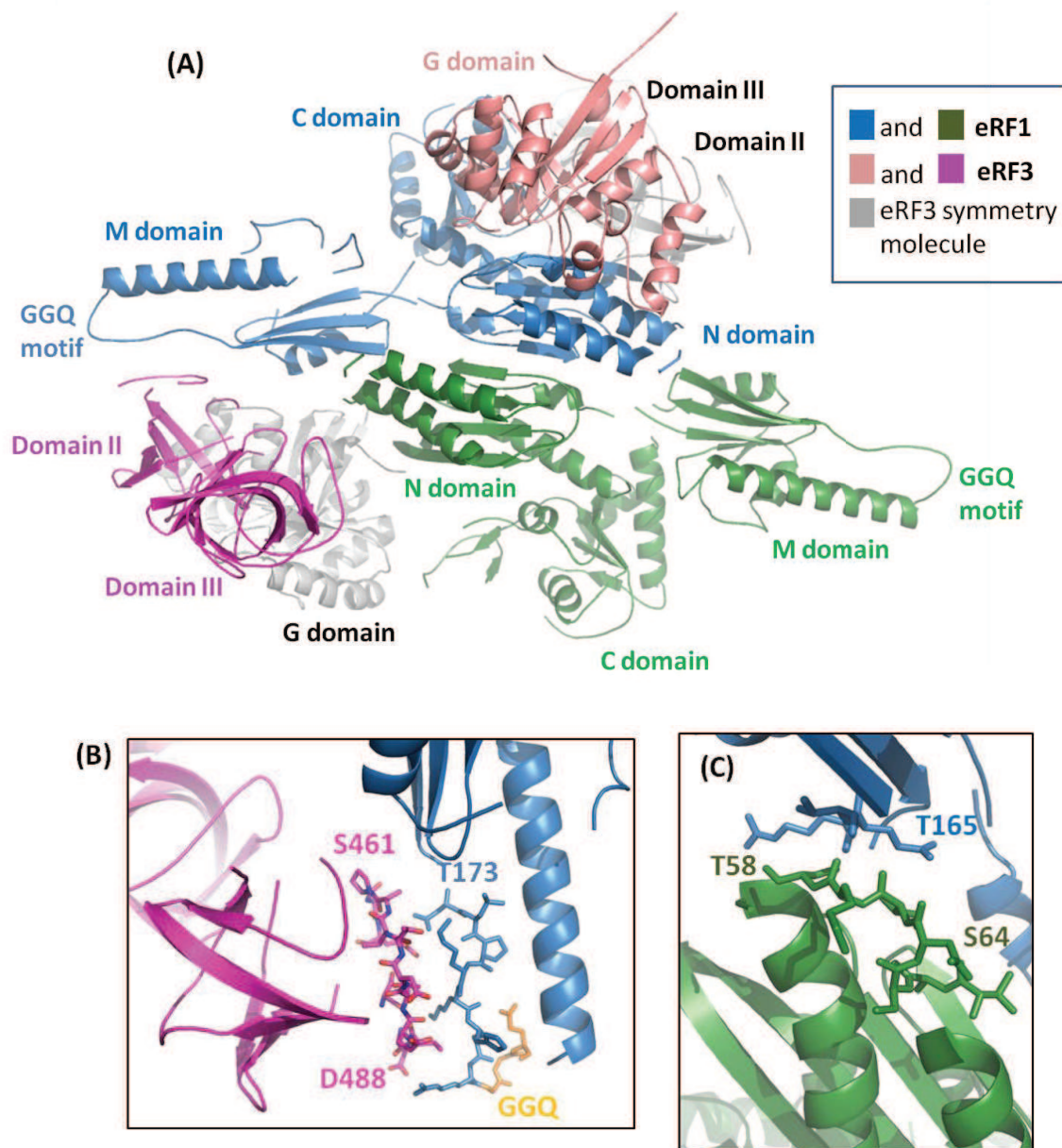
Figure 46: eRF1-eRF3 crystal structure. The components of the asymmetric unit are marked in green- form I eRF1, blue – form II eRF1, magenta – Domain II and Domain III of eRF3 and salmon – G domain of eRF3. In grey are the components from the symmetry molecule. (B) eRF3 domain II interactions with M domain, GGQ motif is marked in orange. (C) [58]TASNIKS[64] motif, highlighted in red, of form I eRF1, interacting with β-sheet of M domain in form II eRF1.

Furthermore, the mode of eRF1 dimerization in this structure is similar to the previously reported eRF1 dimer (Song et al., 2000) with N domain primarily involved in dimerization interactions (Cheng et al., 2009). Four stranded β-sheet present in each eRF1 N domain, contiguously form a long eight-stranded β-sheet with the innermost β-strand (residue 121-128) contacting the same strand on the other eRF1 in opposite direction. Besides the positioning of the NIKS motif, involved in stop codon recognition, of two eRF1 molecules is particularly interesting as these motifs of both molecules interact differently. The eRF1 form I [59]TASNIKS[65] motif interacts with the M domain of form II thereby creating new contacts and probably leading to the conformational change as discussed below. The eRF1 form II NIKS motif does not interact with M domain and thus, this molecule is present in a relatively closed conformation. These two eRF1 shared 2179 Å^2 area of contact, signifying a strong dimer interface. On the other hand form II eRF1 and eRF3 G-domain surface area was determined to be 5193 Å^2, indicating strong interactions which explains dissociation of eRF1 dimer in solution and eRF1-eRF3 heterodimer formation. In addition, the domain II of eRF3 and M-domain of eRF1 have only 803 Å^2 contact surface representing only weak van der Waals and hydrophobic contacts as have been reported in the previous crystal structure (Cheng et al., 2009).

During molecular replacement, finding the position and orientation of the second eRF1 was difficult, due to its different conformations and flexibility of M and C domains. In the α-β sandwich fold of the C domain most of the β-sheet residues 330-372 were poorly ordered in form II and the electron density for minidomain could not be assigned here. Upon aligning the N domains of both eRF1 molecules it was observed that the M and C domains of form I have moved towards each other, in other words, away from the N domain. This could be because the N domain of form I is exclusively involved in dimerization, while N domain of form II also interacts with eRF3 G domain. Thus, dual interactions make N and C domains more compact in form II eRF1, while in form I, the M and C domains are relatively more mobile and take up an open conformation. In addition, this is aided by the fact that had these molecules been in the same conformation, eRF3 would not be able to bind eRF1 due to steric clashes between the G domain and N terminus of eRF1. Moreover, we noticed that the eRF3 interactions with the two eRF1 moieties at their M domain were remarkably different (Fig. 47). For form I eRF1 molecule, the M-

domain GGQ end interacts with domain II and G domain, while the M domain of the form II eRF1 does not; instead it interacts with domain III and extends its GGQ loop in between the two β-barrels. Probably this is a consequence of eRF1 adopting two different conformations.
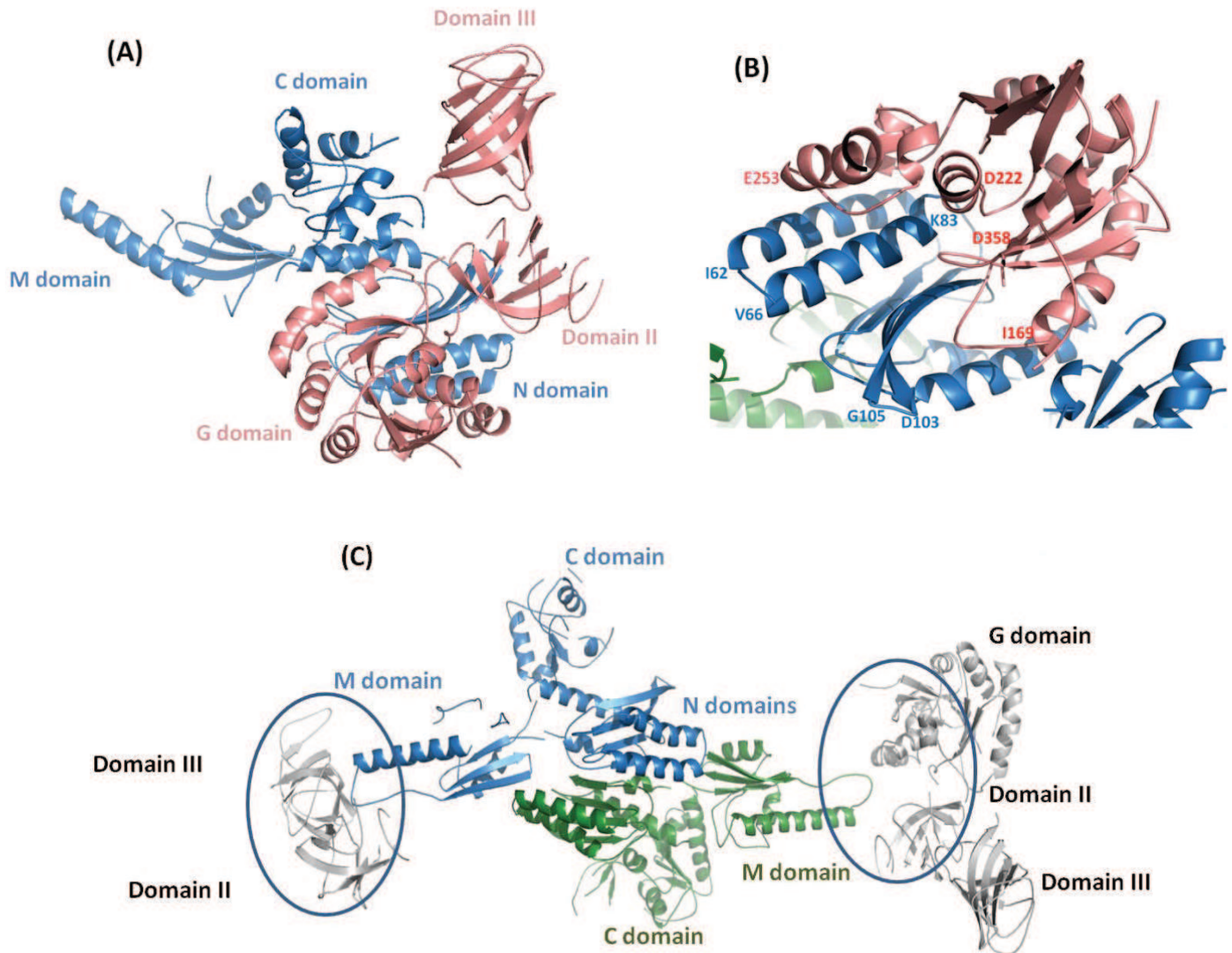
Figure 47: Interactions within the eRF1-eRF3 protein complex. (A) Form II eRF1 (blue) and G domain eRF3 (salmon) interactions. (B) Zoomed-in view for eRF1 and G domain interactions, some residues are marked. (C) Different GGQ and M-domain interactions of the two eRF1molecules (blue and green) with symmetry-related eRF3 molecules (grey). For simplicity, eRF3 from the asymmetric unit is not shown.

Earlier, the eRF1 structure was thought to be a rigid entity (Song et al., 2000), but off late, it has been reported to adopt different conformations depending on its interacting partner, as observed upon binding to the ribosome and to eRF3 (Cheng et al., 2009; des Georges et al., 2014). On comparing the present crystal structure with the previously published eRF1-eRF3 crystal structure, the M-domain of form II, which contains the GGQ

motif and is required for peptide release, was observed to be moved by about 60° (Fig. 48). Surprisingly, the eRF1 conformation observed here is yet different from the conformation of the ternary complex when recruited on ribosome by yet another ~27° shift in the long stem of the M domain (Fig. 48). In spite of the conformational flexibility, "Y"- shaped architecture of eRF1 is retained in both the copies present in the asymmetric unit.
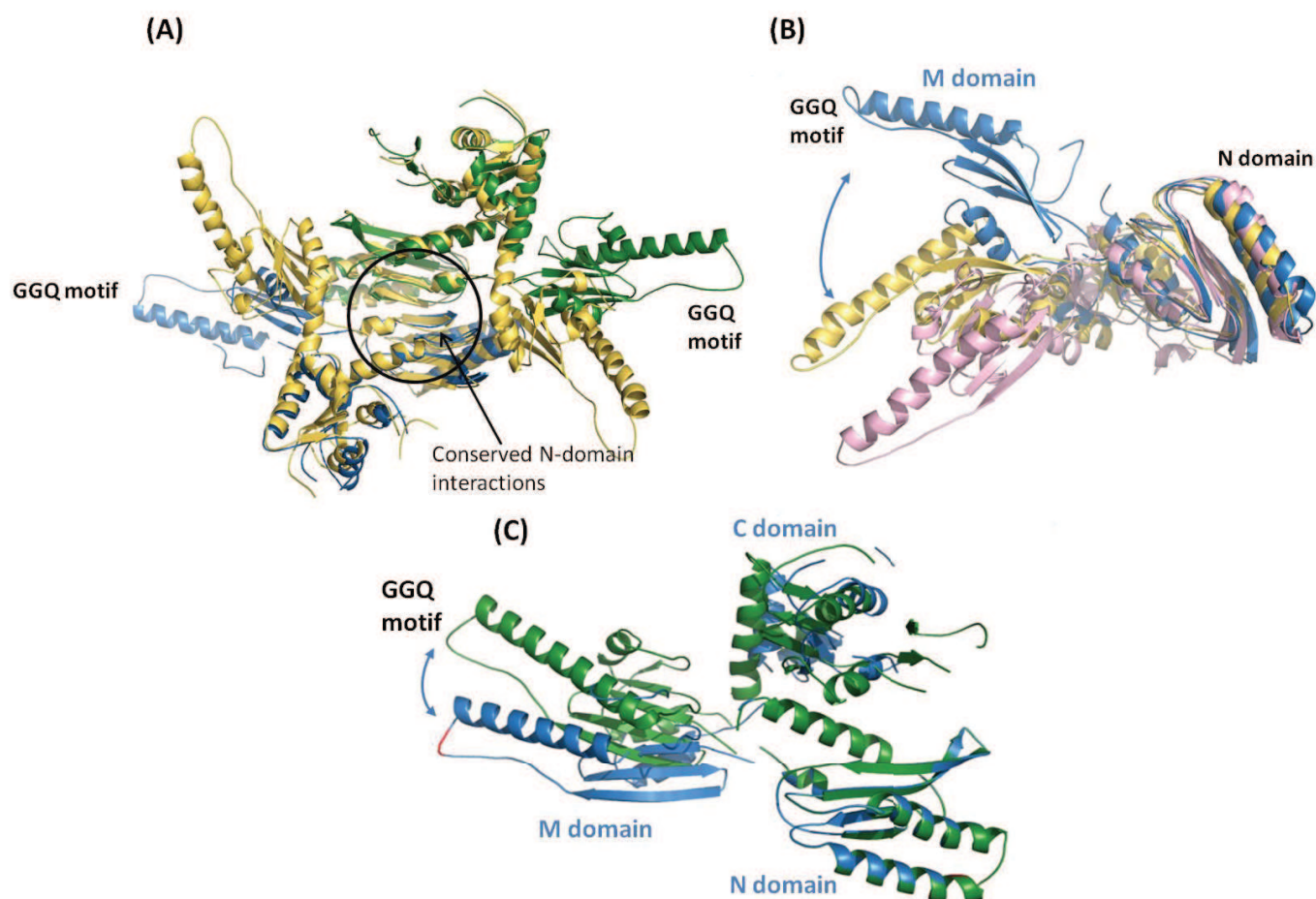
Figure 48: Conformational flexibility in eRF1 structure. Form II eRF1 molecule in blue, N domain superposed on (A) eRF1 from previous crystal structure (PDB code 3E1Y) in yellow, (B) the structure of eRF1 when bound to the ribosome (PDB code 3J5Y) in pink, (C) Form I eRF1 molecule in green, N domain. For simplicity, eRF3 molecule is not shown.

Thus, as explained above, eRF1-eRF3 domain interactions observed in this structure vary from those observed by (Cheng et al., 2009) in the previously reported crystal structure. Most likely the presence of the G domain alters the crystal packing, leading to $P4_12_12$ space group instead of $P4_3$. We performed rigid body refinement and further refinement is ongoing. Due to limitation in resolution of the data, we cannot localise many side chains. Nevertheless backbone residues located within 5 Å vicinity of eRF1 are

considered to be involved in eRF1-eRF3 interactions which are listed in Table 12. A further comparison of this structure with the eRF1-eRF3-GMPPCP ternary complex bound on the ribosome is present in the next section, which highlights the M domain flexibility of eRF1.

Table 12: Interacting residues in eRF1-eRF3 structure. * indicates interactions with a symmetry-related molecule.

eRF1 residues		eRF3 residues	
M domain (form I)	Ser 154	Glu 562, Asn 549	**Domain III***
	Lys 171	Asn 549, Asp 619, Glu 620	
	Thr 173	Tyr 548	
	Glu 261	Glu 620	
	Tyr 558	Glu 559	
	Asn 262	Glu 560	
C domain (form I)	Asp 297	Asn 433, Glu 211	**G domain***
	Asp 418	Thr 347	
	Gln 420	Thr 347	
N domain (form II)	Lys 16	Arg 371	**G domain**
	Lys 18	Glu 370	
	Lys 19	Glu 370	
	Leu 20	Thr 364	
	Ser 23	Pro 363	
	Leu 24	Thr 364	
	Phe 56	Asn 255	
	Ser 60	Glu 257	
	Arg 68	Asn 255	
	Leu 72, Thr 76	Glu 257	
M domain (form II)	Lys 178	Thr 491, Thr 493	**Domain II**
	Lys 179	Ser 486, Asp 488	

New crystals have been obtained with the same condition and must be tested at the synchrotron for diffraction. A higher resolution data set could provide better insights into the interaction of these two proteins.

3.3 Pre-termination complex

Human 80S ribosomes purified as part of this work were used to form a pre-termination complex composed of the eRF1-eRF3 protein complex with and without nucleotide, mRNA and tRNA-Lys (UUU). The complexes were prepared using direct mixing and incubation of the different components. The property of ribosomes to "breathe", wherein the two subunits are partially dissociated in a manner depending on salt concentration, was exploited. Initial tests were performed with a short mRNA sequence. However, the complexes formed did not have the P-site tRNA present as revealed by cryo-EM reconstructions, instead, only the E-site tRNA was present with a partial occupancy of the release factors on the 80S. The data collection statistics are mentioned in the table 13.

Table 13: Cryo-EM data collection parameters for termination complex.

	Complex with short mRNA	Complex with long mRNA
Complex formation	1: 1.5: 2: 5 80S: mRNA: tRNA: eRF1-eRF3-GMPPCP	1: 10: 2: 5 80S: mRNA: tRNA: eRF1-eRF3
Detector	CMOS Falcon camera 4K X 4K	CMOS falcon camera 4K X 4K
Voltage	300kV	300kV
Pixel size	1.82 Å/pixel	1.14 Å/pixel
Box size	240 X 240 pixels	512 X 512 pixels
Magnification	93k	93k
Total dose	15 ē/Å^2	20 ē/Å^2
Defocus	-0.8 to -3.5 μm	-1 to -5 μm
Number of particles	80,000 particles	120,000 particles

Unexpectedly, in the complex with the short mRNA, a large density was found in the map different from eRF1-eRF3, suggesting the presence of an elongation factor eEF2 bound to ribosome in the factor binding site. A slight change in salt concentration during ribosome purification especially during puromycin treatment, could have led to its retention on the ribosome. This accounted for about 10% of the total data set after particle sorting by image processing (Fig. 49). It was possible to assign this density, unambiguously to eEF2 due to an extra domain as compared to eRF3 and perfect fitting using the eEF2 crystal structure. This complex was found in a ratcheted state unlike the eRF1-eRF3 containing complex (Fig. 50).

Later, a longer mRNA sequence (40 nucleotides) was used to assemble the complex. The absence of P-site tRNA in this complex as well led us to conclude that the breathing of the 2 subunits is not easily feasible in eukaryotic ribosomes, in sharp contrast with prokaryotic ribosomes. Nevertheless, in both complexes, the density for eRF1 as well as eRF3 was correctly identified in the factor binding site on the ribosome (Fig. 50) and is similar to that observed by (des Georges et al., 2014) in the pre-termination complex.
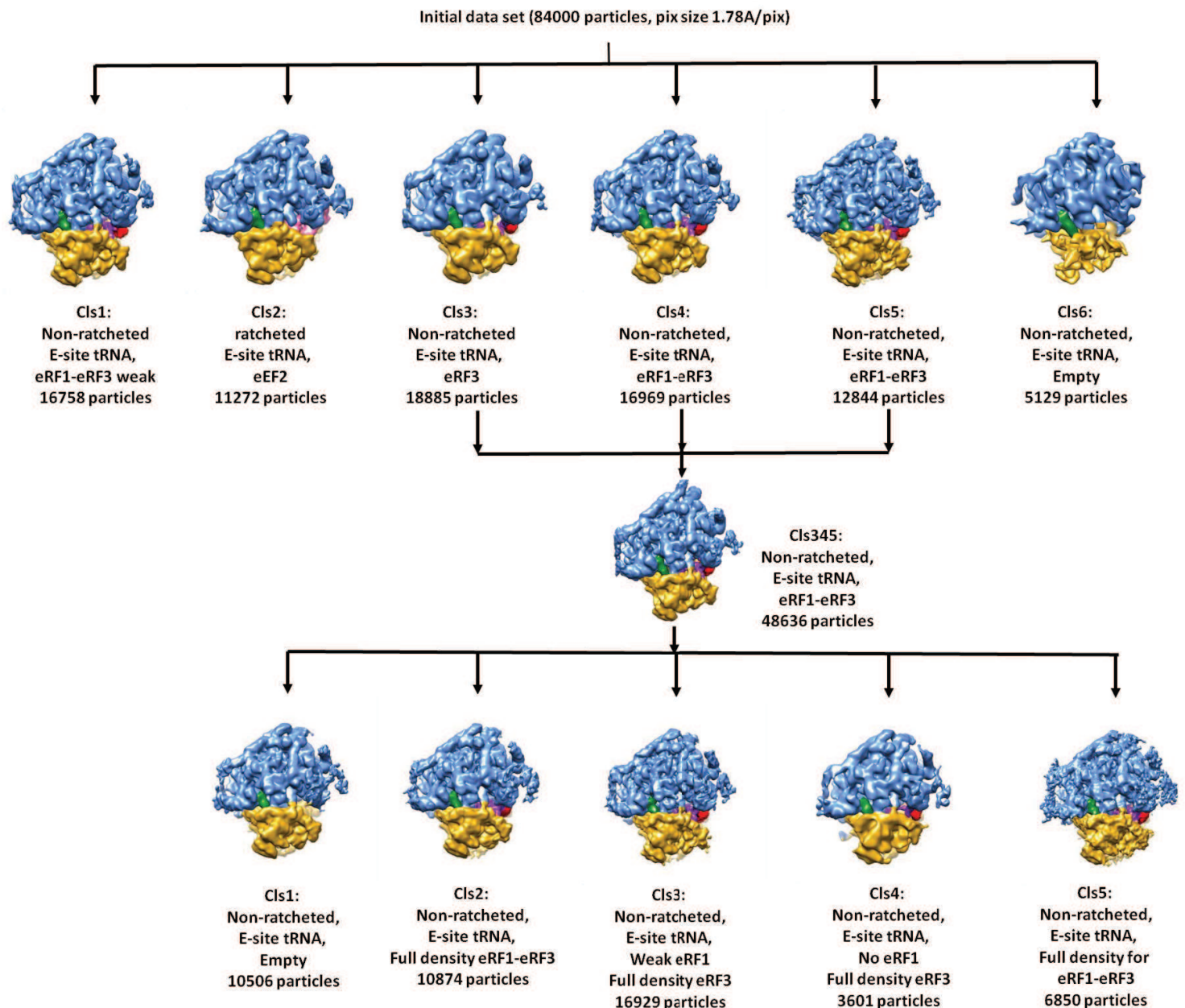
Figure 49: Scheme for splitting the data set with short mRNA. A similar scheme was used for splitting the data set with long mRNA.

Remarkably, eRF1 exists as a dimer in solution but in the presence of eRF3 it dissociates, probably due to relatively weak interactions. This allows formation of eRF1-eRF3 functional heterodimer as observed on the native gel, and as detected in the termination complex. Thus, it is the very presence of its functional partner that allows eRF1 to go from a non-functional homodimer to a heterodimer form. But the eRF1 M domain is not accommodated exactly in the PTC in neither our complex, nor those reported before. Thus, it requires GTP hydrolysis to accommodate in the PTC and assist in peptide release.
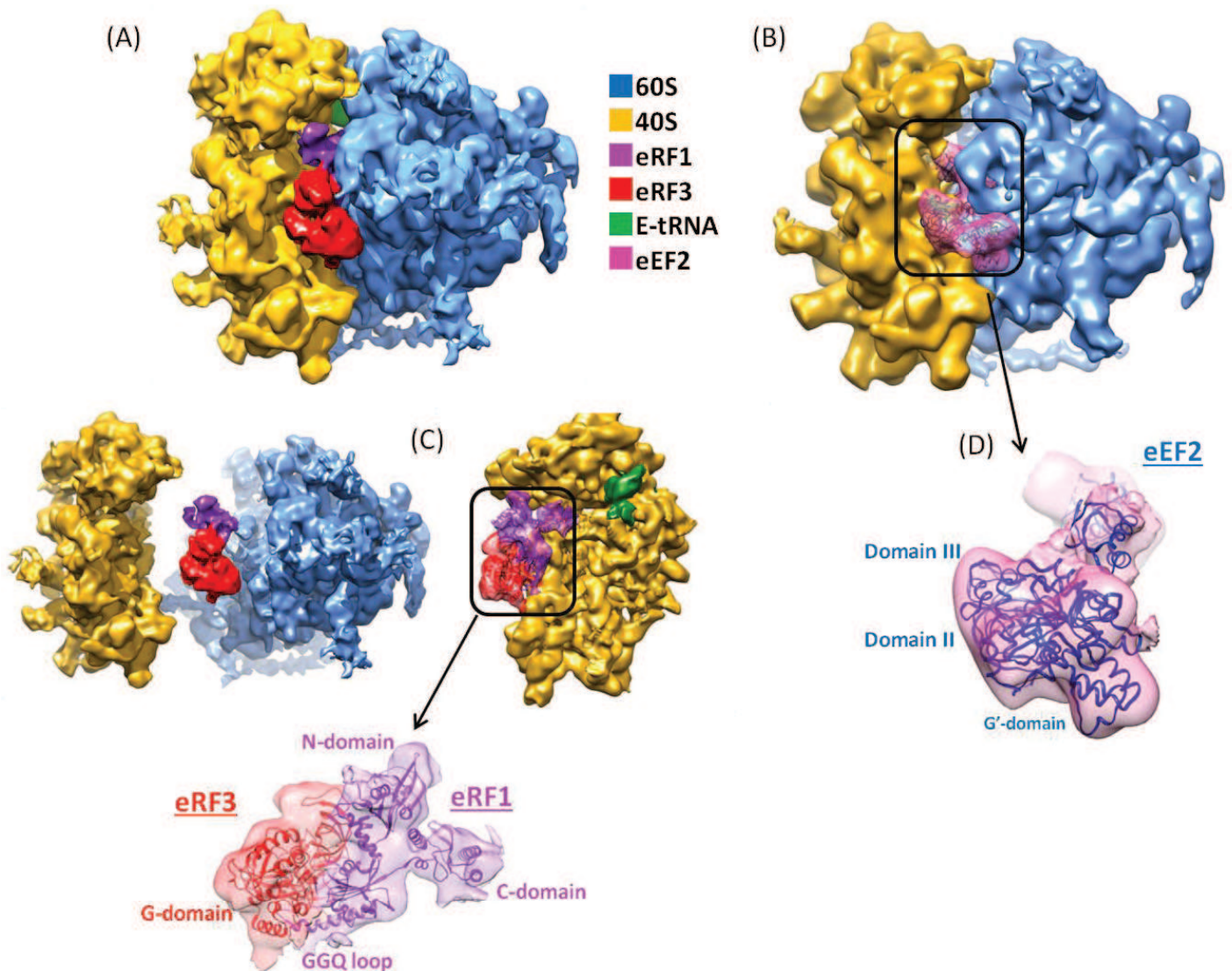
Figure 50: The termination complex reconstituted *in-vitro*, with short mRNA sequence. (A) Side view of the complex showing the release factor at the factor binding site, eRF1 in purple, eRF3 in red and E-site tRNA in green. (B) Side view of the complex showing eEF2. (C) The densities split corresponding to the two subunit and the release factors. The fitted eRF1-eRF3 complex, PDB code 3J5Y shows that the GGQ motif is pointing away from the PTC. (D) The density corresponding to eEF2 with the fitted crystal structure, PDB code 3DNY. Large and small subunits are coloured blue and yellow respectively.

Interestingly, on aligning the crystal structure of eRF1-eRF3 with the density observed on ribosome, the M domain harbouring the functional GGQ motif of eRF1, responsible for peptide release, was found to be in a different conformation. It has moved by about 90° towards the PTC. The N domains were superimposed, but the C domain is also partially moved. eRF3 interactions in the crystal structure described above are such that domain 2 and 3 are in close vicinity of M domain while G domain interacts majorly with N and C domains of eRF1. Such interactions have not been reported so far, and may represent crystal packing requirements rather than *in-vivo* interactions.

Due to heterogeneity in the sample, the collected data sets in both cases were sorted out into different 3D classes using the Relion software (Scheres, 2012). In case of the short mRNA, the eRF1-eRF3 protein complex was present in all the classes while for the complex with long mRNA, only one class out of the six had the eRF1-eRF3 protein complex. This decreased factor occupancy could be due to the absence of a nucleotide analogue GMPPCP in the eRF1-eRF3 complex, but this needs to be analysed further. The presence of nucleotides can help stabilise the factor binding to the ribosome, and its absence could make the sample heterogeneous and can even lead to factor dissociation.

We have recently collected data for termination complex with long mRNA on the in-house Titan Krios, which needs to be processed and could provide high resolution structural insights into eRF1-eRF3 binding on the ribosome as observed for the reference 80S structure detailed in section 3.1.1.

3.4 HIV-1 Gag interactions with ribosome

HIV-1 Gag poly protein has been known to regulate its own translation (Anderson and Lever, 2006). However, its mechanism of action remains unclear. Previous studies using mass spectrometry, and our collaborators narrowed down several ribosomal proteins that interact with Gag (Jager et al., 2012). Of these, RPL30 (old name L7) and RPL24 (old name L26) were verified using the yeast two-hybrid system, the sucrose density gradient analysis and other biochemical assays like co-immunoprecipitation and FLIM-FRET (de Rocquigny et al., 2014).

However, most of these studies are centred on Gag's interaction with soluble ribosomal proteins (Beyer et al., 2013). Here, we tried to investigate Gag's interaction with the above mentioned proteins, in context of the whole ribosome. We performed polysome profile analysis on HeLa cells transfected with plasmids carrying the Gag insert. The sucrose fractions were analysed by western blot analysis for the simultaneous presence of Gag and RPL30. Gag was observed to be co-sedimented with polysomes and less with monosomes (Fig. 51). Moreover, to rule out non-specific interactions, negative controls were performed using cells transfected with EGFP and their polysome profiles were analysed (Fig. 51).
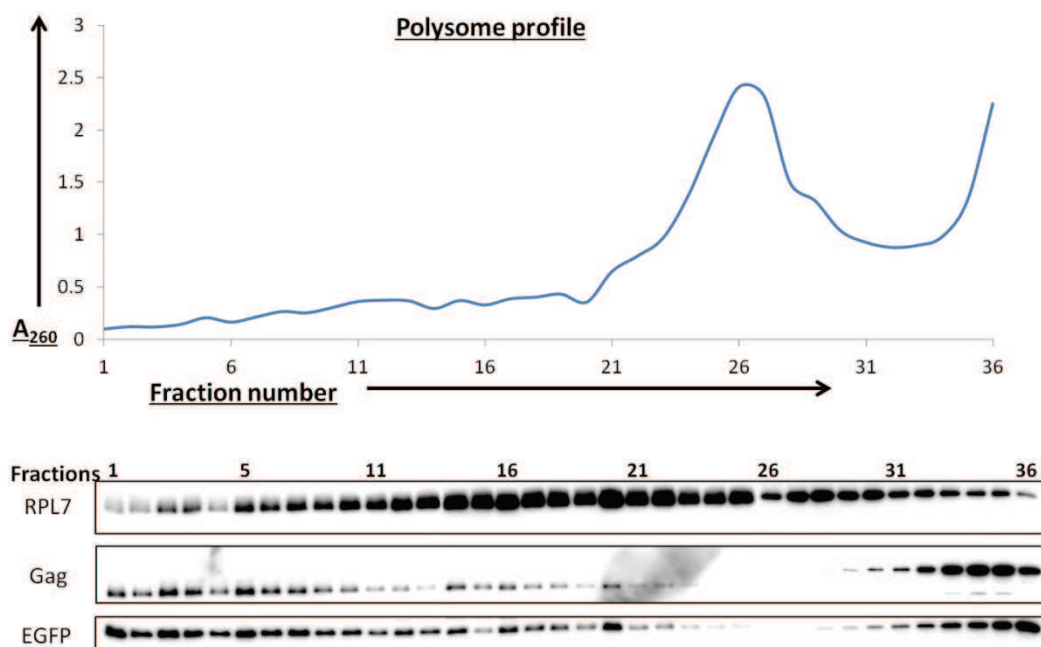
Figure 51: Polysome profile for Gag and EGFP co-transfected cells. Below is the western blot analysis for each of the fractions, detecting RPL7, Gag and EGFP.

These control experiments would help in differentiating *gag* being translated using its own mRNA, from the ribosome-bound gag polyprotein acting as a factor for regulating its own translation. Strangely, EFGP profiles were similar to gag polysome profiles, so we performed another set of controls. Density gradient profiles were checked for Gag and EGFP co-transfected HeLa cells after EDTA treatment which would allow dissociation of the two ribosomal subunits (Nolan and Arnstein, 1969). However, even with this approach, interaction of Gag with either of the subunits remains inconclusive. Thus, future studies are needed to address this in more detail.
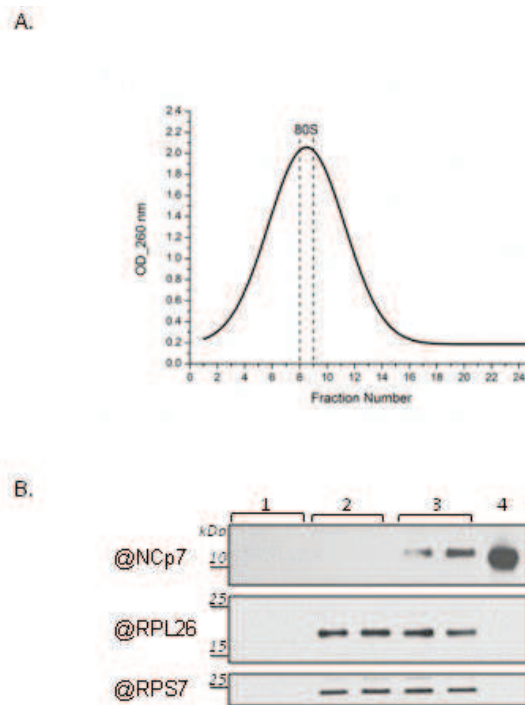
Figure 52: NCp7 co-sedimentation with 80S ribosomes. (A) Sucrose gradient fractionation profile of purified 80S ribosomes incubated with NCp7. (B) Western blot analysis of fractions 8-9 from sucrose gradient with only NCp7 peptide (lanes 1); only 80S ribosomes (lanes 2); 80S ribosomes and NCp7, together (lanes 3). As a control, 90 nM of NCp7 peptide was loaded on the SDS-gel (4).

Interestingly, despite our inconclusive results related to Gag interaction with RPL30 as an entity of ribosome, it was determined that NCp7 region of Gag was crucial for its interaction with RPL30. To this, we used human 80S ribosomes purified from HeLa cells (Khatter et al., 2014) and chemically synthesised NCp7 peptide, to check for co-sedimentation on sucrose density gradient. The association of NCp7 with ribosomes was safely inferred due to co-localisation of both these entities as determined by western blot analysis (Fig. 52) (Anton et al., manuscript submitted).

4.   DISCUSSION

The human ribosome structure had been a relatively unexplored territory until a year ago (2013). One of the reasons for that is the complexity of the human ribosome in terms of composition, long RNA expansion segments at its surface and its instability when purified from human cells, parameters that all favoured the belief that the human 80S was not amendable for high-resolution structure investigation. Indeed, for decades the very first requirement of pure, homogenous sample was not fulfilled for structural studies. This signified a major constraint in studying eukaryotic translational aspects. In order to investigate these areas, I worked on the purification and characterization of human ribosomes purified from HeLa cells. In this thesis, I have developed and detailed a method to purify homogenous 80S ribosomes and the two subunits, 60S and 40S, from HeLa cells. These ribosomes were characterized biophysically and using structural techniques for their purity and uniformity. A crucial aspect for this study was the usage of cryo-EM as an analytical tool rather than exclusively to obtain 3D reconstructions. This allowed visual control of samples prepared under various conditions and thus, helped in standardizing the purification protocol. And eventually, obtaining crystals of human 80S in capillaries since only samples with an even distribution on cryo-EM grids were used for crystallisation. These crystals were reproduced in sitting drops and diffracted X-rays up to 26 Å. Despite the low resolution diffraction of these crystals, they signify a major breakthrough in human 80S structural studies and pave the way for future high resolution work (Khatter et al., 2014).

Eukaryotic translation itself is a relatively less investigated target with therapeutic potential for fungal and viral infections as well as cancer. Also, the growing problem of antibiotic side-effects needs to be examined. These ribosomes can prove to be essential to understand and tackle such issues. In the long term these studies will be useful for drug discovery, targeting the pathogen more specifically rather than inducing interactions with human ribosomes.

As the old adage goes, the more we learn, the more we realise how little we know. In spite of more than a decade's research on translation mechanisms, our understanding of eukaryotic translation is far from complete, especially with regard to the regulatory mechanisms. The purified 80S ribosomes fulfil an essential criterion for studying these mechanisms in that they provide the core complex to which various ribosomal factors bind during the initiation, elongation and termination phases of translation. We tried to

understand the principle of translation termination using these purified ribosomes. The tricky part was to assemble the termination complex *in vitro* using separately purified and assembled class I (eRF1) and II (eRF3) termination factors. I assembled the 80S complex with these factors, mRNA and tRNA by varying temperature and salt concentrations. In the cryo-EM structures reported here, even though the eRF1-eRF3 protein complex is bound in the factor binding site and GTPase centre, the absence of P-site tRNA makes this complex not exactly a termination complex. Instead, an E-site tRNA is bound, which could be either the co-purified tRNA (as for empty 80S ribosomes, Khatter et al., 2014) or the added tRNA-Lys which was not positioned correctly.

Consequently, mRNA binding and positioning was the key problem that we encountered. First, a short mRNA sequence was used which probably did not have enough residues to allow mRNA binding. Later, even a longer mRNA sequence didn't help tRNA binding into the P-site as revealed by cryo-EM reconstruction. It is possible that the *in vitro* binding of the eRF1-eRF3 factor complex to the ribosome is not influenced by the presence of mRNA or the tRNA being in the correct reading frame. This needs to be verified and data recently collected (7.2014) on the Titan Krios cryo electron microscope could provide high resolution details of the termination complex. In principle it should become feasible to analyse atomic details of release factor interactions with the ribosome with the new equipment and approach for data analysis as explained earlier for the reference 80S complex, provided that the functional complex has formed under the conditions we currently used.

In the context of translation termination, another aspect that I tried to understand was the full length eRF1-eRF3 protein complex. Both these release factors have been studied individually as crystal structures but the full-length eRF3 structure is not yet known. I purified both proteins and obtained well-diffracting crystals for the complex. However, initial trials led to crystals in conditions where the complex was dissociating. On modifying the final buffer composition of the complex, I could obtain crystals with both proteins present, which diffracted to about 4 Å. Interestingly, upon structure determination by molecular replacement eRF1 was observed to be in a new conformation that has not been reported before. While interactions between the M domain of eRF1 and C terminal domain of eRF3 interactions are similar to those observed before, there are also unpredicted

interactions between the G-domain of eRF3 and the N-domain of eRF1. The structure refinement is still ongoing for this complex and I have more crystals with this condition with nucleotide analogues bound (co-crystallised or soaked) that need to be tested at the synchrotron. Determining this structure to high resolution with the new crystals and even performing cryo-EM studies on eRF1-eRF3 protein complex in solution will shed light on the domain movements and flexibility of the eRF1 molecule within the eRF1-eRF3 complex. Furthermore, these termination complex studies fundamentally aim at answering the mechanism of release factor binding and decoding with the NIKS motif interaction on the ribosome to permit peptide release with help of the GGQ motif. To elucidate these interactions, high-resolution structures must be obtained in the context of the 80S ribosome.

Lastly, the purified ribosomes can also be used for studying protein or factor interactions with 80S or ribosomal subunits, as presented here for the HIV-1 Gag polyprotein. This work was performed in collaboration with Yves Mély's group at the Faculty of Pharmacy, Illkirch. They detected that the NCp7 region of Gag was indispensable for interactions with ribosomal protein RPL30. We verified these interactions in the context of full 80S, and observed co-sedimentation of 80S and NCp7. This binding of Gag to 80S could be a mechanism of Gag regulating its own translation, but this needs to be verified with further experiments.

In summary, the human ribosome continues to be an enigma not just structurally but even in terms of functional characteristics. Recent studies have confirmed the structural conservation of the ribosomal core during evolution from single-cell prokaryotes to the complex multi-cellular eukaryotes, with eukaryotic specific elements being added on the solvent-exposed area. We report here a human 80S structure at about 5 Å resolution in which most secondary structure of the rRNA as well as those of protein α-helices and some ß-strands can be distinguished. Further refinement of this structure is under process and could provide near-atomic details about this huge translation machinery. Finally in terms of function, the analysis of the termination phase needs to be pursued in greater detail preferably using an *in vitro* system, to unveil the molecular mechanism of stop codon recognition by eRF1 in the ribosomal decoding centre of the 40S, and peptide release from the tRNA in the peptidyl transferase centre of the 60S ribosomal subunit.

5. Appendix

5.1 Plasmid purification

Macherey Nagel (Mini or Midi preparation from 5ml or 20 ml culture respectively)

1. Cultivate and harvest bacterial cells from a culture grown overnight to saturation. Pellet cells in a standard benchtop microcentrifuge for 30 seconds at 11,000 x g. Discard the supernatant.

2. For cell lysis add 250μl buffer A1 and resuspend the cell pellet completely by vortexing or pipetting up and down. Add 250 μl buffer A2 and mix gently by inverting the tube 6–8 times. Do not vortex to avoid shearing of genomic DNA. Incubate at room temperature for up to 5 min or until lysate appears clear. Add 300μl buffer A3 and mix thoroughly by inverting the tube 6–8 times. Do not vortex to avoid shearing of genomic DNA.

3. Clarification of lysate by centrifugation for 5 min at 11,000 x g at room temperature.

4. To bind DNA, place a NucleoSpin® Plasmid column in a Collection Tube (2 mL) and pipette 750 μl of the supernatant onto the column. Centrifuge for 1 min at 11,000 x g. Discard flow-through and place the NucleoSpin® Plasmid column back into the collection tube.

5. Add 600 μL Buffer A4 (supplemented with ethanol) and centrifuge for 1 min at 11,000 x g. Discard flow-through and place the NucleoSpin® Plasmid column back into the empty collection tube.

6. Centrifuge for 2 min at 11,000 x g and discard the collection tube.

7. Elute DNA by placing the NucleoSpin® Plasmid column in a 1.5 mL microcentrifuge tube and add 50 μL Buffer AE/ water. Incubate for 1 min at room temperature. Centrifuge for 1 min at 11,000 x g.

5.2 Expression and purification of TEV Protease

Transformation of competent cells:

- Incubate plasmid with E. Coli BL21(DE3) on ice for 30 minutes, followed by a heat shock at 42 °C for 45 seconds.

- Immediately cool on ice for 2 minutes and incubate with LB medium at 37 °C for 1 hour.

- After incubation, spread the culture on LB Agar plate containing Kanamycin for overnight growth.

Culture and protein expression

- Inoculate a single colony from the agar plate in 100 ml of LB containing Kanamycin for overnight growth at 37 °C, with constant stirring.

- Inoculate 6L LB media with equal volumes of the overnight cultivated culture for 16 hours.

- Pellet the cells by centrifugation for 20min at 4000rpm.

Cell Lysis and purification

- Resuspend the cells in lysis buffer (20mM HEPES pH 8 , 1M KCl , 30mM Imidazole, 0.1 % Triton X-100, 1mM βME).

- Sonicate for 4 minutes (twice) on ice with 40% amplitude.

- Centrifuge the lysate at 300000g for 45' at 4 °C to recover the supernatant containing the protein.

- Incubate the supernatant with Ni-NTA resin with constant stirring for 2 hours at 4°C.

- Centrifuge at 1000g, 2' at 4 °C and load the resin mixture onto a small disposable column.

- Wash twice with lysis buffer and elute with 20x (1 ml) of elution buffer (20mM HEPES pH8, 1M KCl, 30mM Imidazole, 0.1 % Triton X-100, 1mM βME, 50 mM EDTA). Collect each fraction separately and run on 15% SDS-PAGE.

Dialysis and Freezing

- Pool fractions containing the TEV protein and dialyse overnight against dialysis buffer (50mM Tris pH 8, 50mM NaCl, 1mM DTT, 0.1% Triton X-100, 1mM EDTA).

- Concentrate the dialysed protein using Amicon filter 10kDa cutoff.

- Add 50% autoclaved glycerol. Aliquot the purified TEV in 100µl fractions and freeze in liquid nitrogen. The protein must be stored in -80 °C, for upto 2 years.

5.3 Autoinducible media

STOCK SOLUTIONS			
1M MgSO4	50x 5052	20x NPS	ZY
	110 ml H_2O	270 ml H_2O	925 mL H2O
	37.5g glycerol	19.8g $(NH_4)_2SO_4$	10g tryptone
	3.75g D-(+)-glucose	40.8g KH_2PO_4	5g Yeast Extract
	15g Δ-lactose	42.6g Na_2HPO_4	
Autoclave all solutions (except 50X 5052)			
For 1L of culture : - 928ml ZY - 1ml of 1M MgSO4 - 20ml of 50x 5052 - 50ml of 20x NPS			

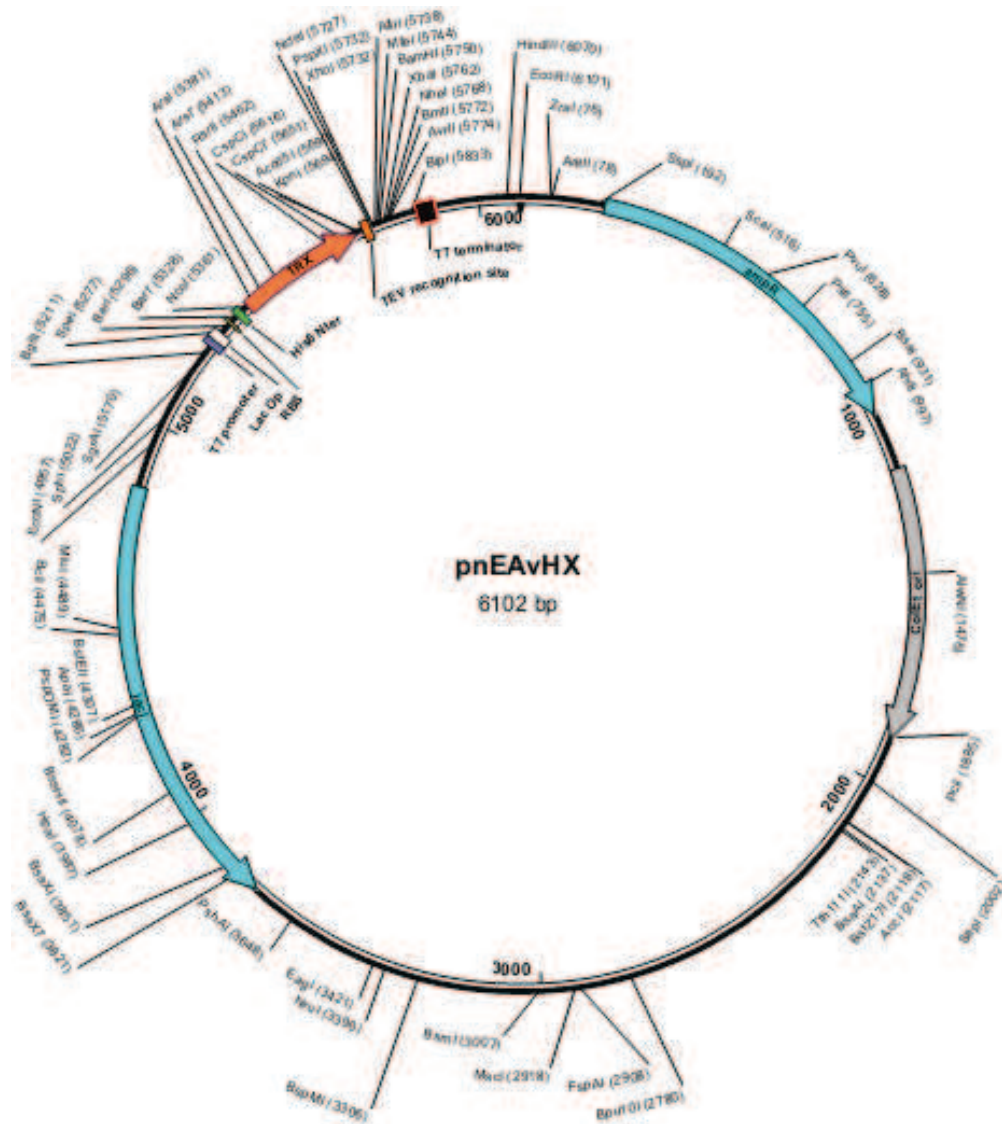## 5.4 Plasmid pnEAvHX vector map

Figure 53: pnEAvHX vector used for eRF3 cloning in *E. Coli* DH5α bacteria.

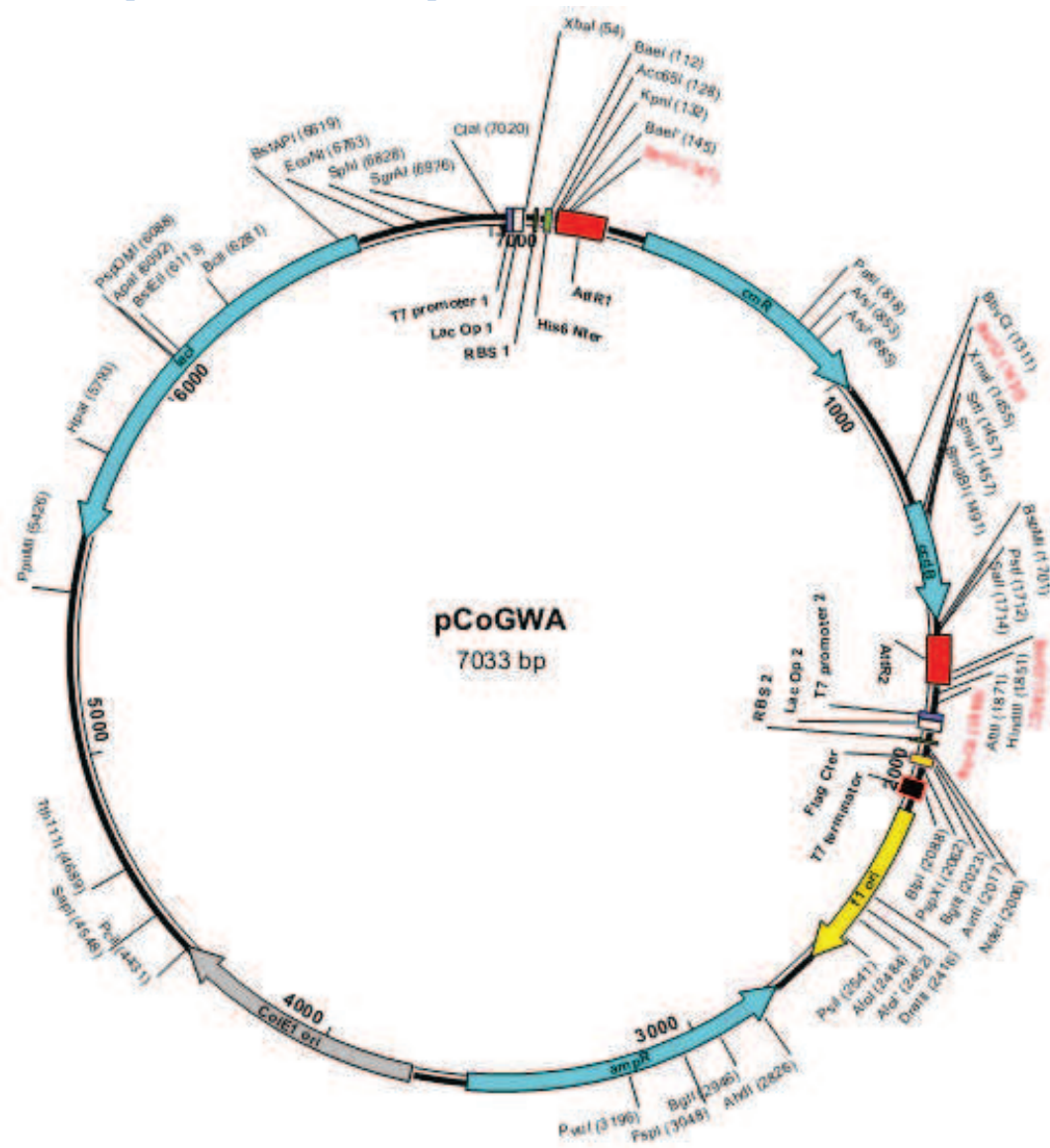## 5.5 Plasmid pCoGWA vector map

Figure 54: pCoGWA vector used for eRF1 cloning in *E. Coli* DH5α bacteria.

6. References

Alkalaeva, E.Z., Pisarev, A.V., Frolova, L.Y., Kisselev, L.L., and Pestova, T.V. (2006). In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. Cell *125*, 1125-1136.

Allen, G.S., Zavialov, A., Gursky, R., Ehrenberg, M., and Frank, J. (2005). The cryo-EM structure of a translation initiation complex from Escherichia coli. Cell *121*, 703-712.

Amunts, A., Brown, A., Bai, X.C., Llacer, J.L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S.H., and Ramakrishnan, V. (2014). Structure of the yeast mitochondrial large ribosomal subunit. Science *343*, 1485-1489.

Anderson, E.C., and Lever, A.M. (2006). Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation. Journal of virology *80*, 10478-10486.

Andjelkovic, N., Zolnierowicz, S., Van Hoof, C., Goris, J., and Hemmings, B.A. (1996). The catalytic subunit of protein phosphatase 2A associates with the translation termination factor eRF1. The EMBO journal *15*, 7156-7167.

Anger, A.M., Armache, J.P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. Nature *497*, 80-85.

Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S*., et al.* (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. Proceedings of the National Academy of Sciences of the United States of America *107*, 19748-19753.

Bacharach, E., Gonsky, J., Alin, K., Orlova, M., and Goff, S.P. (2000). The carboxy-terminal fragment of nucleolin interacts with the nucleocapsid domain of retroviral gag proteins and inhibits virion assembly. Journal of virology *74*, 11027-11039.

Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassucci, R.A., Sweet, R., Frank, J., Moore, P.B., and Steitz, T.A. (1998). A 9 A resolution X-ray crystallographic map of the large ribosomal subunit. Cell *93*, 1105-1115.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science *289*, 905-920.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta crystallographica Section D, Biological crystallography *67*, 271-281.

Becker, T., Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Sieber, H., Motaal, B.A., Mielke, T., Berninghausen, O., and Beckmann, R. (2011). Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. Nature structural & molecular biology *18*, 715-720.

Becker, T., Bhushan, S., Jarasch, A., Armache, J.P., Funes, S., Jossinet, F., Gumbart, J., Mielke, T., Berninghausen, O., Schulten, K*., et al.* (2009). Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. Science *326*, 1369-1373.

Becker, T., Franckenberg, S., Wickles, S., Shoemaker, C.J., Anger, A.M., Armache, J.P., Sieber, H., Ungewickell, C., Berninghausen, O., Daberkow, I*., et al.* (2012). Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. Nature *482*, 501-506.

Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001a). Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell *107*, 361-372.

Beckmann, R., Spahn, C.M., Frank, J., and Blobel, G. (2001b). The active 80S ribosome-Sec61 complex. Cold Spring Harbor symposia on quantitative biology *66*, 543-554.

Belin, S., Hacot, S., Daudignon, L., Therizols, G., Pourpe, S., Mertani, H.C., Rosa-Calatrava, M., and Diaz, J.J. (2010). Purification of ribosomes from human cell lines. Current protocols in cell biology / editorial board, Juan S Bonifacino [et al] *Chapter 3*, Unit 3 40.

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 A resolution. Science *334*, 1524-1529.

Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. Science *330*, 1203-1209.

Berk, V., Zhang, W., Pai, R.D., and Cate, J.H. (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*, 15830-15834.

Berndt, U., Oellerer, S., Zhang, Y., Johnson, A.E., and Rospert, S. (2009). A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. Proceedings of the National Academy of Sciences of the United States of America *106*, 1398-1403.

Bertram, G., Bell, H.A., Ritchie, D.W., Fullerton, G., and Stansfield, I. (2000). Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. RNA *6*, 1236-1247.

Beyer, A.R., Bann, D.V., Rice, B., Pultz, I.S., Kane, M., Goff, S.P., Golovkina, T.V., and Parent, L.J. (2013). Nucleolar trafficking of the mouse mammary tumor virus gag protein induced by interaction with ribosomal protein L9. Journal of virology *87*, 1069-1082.

Bhushan, S., Meyer, H., Starosta, A.L., Becker, T., Mielke, T., Berninghausen, O., Sattler, M., Wilson, D.N., and Beckmann, R. (2010). Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. Molecular cell *40*, 138-146.

Blobel, G., and Sabatini, D. (1971). Dissociation of mammalian polyribosomes into subunits by puromycin. Proceedings of the National Academy of Sciences of the United States of America *68*, 390-394.

Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J.D., and Stark, H. (2005). Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. Structure *13*, 1695-1706.

Brenner, S., Barnett, L., Katz, E.R., and Crick, F.H. (1967). UGA: a third nonsense triplet in the genetic code. Nature *213*, 449-450.

Brenner, S., Stretton, A.O., and Kaplan, S. (1965). Genetic code: the 'nonsense' triplets for chain termination and their suppression. Nature *206*, 994-998.

Briggs, J.A., Riches, J.D., Glass, B., Bartonova, V., Zanetti, G., and Krausslich, H.G. (2009). Structure and assembly of immature HIV. Proceedings of the National Academy of Sciences of the United States of America *106*, 11090-11095.

Brown, C.M., and Tate, W.P. (1994). Direct recognition of mRNA stop signals by Escherichia coli polypeptide chain release factor two. The Journal of biological chemistry *269*, 33164-33170.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M.*, et al.* (2002). The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC bioinformatics *3*, 2.

Carrasco, N., Hiller, D.A., and Strobel, S.A. (2011). Minimal transition state charge stabilization of the oxyanion during peptide bond formation by the ribosome. Biochemistry *50*, 10491-10498.

Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N., and Noller, H.F. (1999). X-ray crystal structures of 70S ribosome functional complexes. Science *285*, 2095-2104.

Chandramouli, P., Topf, M., Menetret, J.F., Eswar, N., Cannone, J.J., Gutell, R.R., Sali, A., and Akey, C.W. (2008). Structure of the mammalian 80S ribosome at 8.7 A resolution. Structure *16*, 535-548.

Chatel-Chaix, L., Boulay, K., Mouland, A.J., and Desgroseillers, L. (2008). The host protein Staufen1 interacts with the Pr55Gag zinc fingers and regulates HIV-1 assembly via its N-terminus. Retrovirology *5*, 41.

Chavatte, L., Seit-Nebi, A., Dubovaya, V., and Favre, A. (2002). The invariant uridine of stop codons contacts the conserved NIKSR loop of human eRF1 in the ribosome. The EMBO journal *21*, 5302-5311.

Chen, J., Petrov, A., Tsai, A., O'Leary, S.E., and Puglisi, J.D. (2013). Coordinated conformational and compositional dynamics drive ribosome translocation. Nature structural & molecular biology *20*, 718-727.

Cheng, Z., Saito, K., Pisarev, A.V., Wada, M., Pisareva, V.P., Pestova, T.V., Gajda, M., Round, A., Kong, C., Lim, M.*, et al.* (2009). Structural insights into eRF3 and stop codon recognition by eRF1. Genes & development *23*, 1106-1118.

Cimarelli, A., and Luban, J. (1999). Translation elongation factor 1-alpha interacts specifically with the human immunodeficiency virus type 1 Gag polyprotein. Journal of virology *73*, 5388-5401.

Clemons, W.M., Jr., Brodersen, D.E., McCutcheon, J.P., May, J.L., Carter, A.P., Morgan-Warren, R.J., Wimberly, B.T., and Ramakrishnan, V. (2001). Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: purification, crystallization and structure determination. Journal of molecular biology *310*, 827-843.

Connell, S.R., Takemoto, C., Wilson, D.N., Wang, H., Murayama, K., Terada, T., Shirouzu, M., Rost, M., Schuler, M., Giesebrecht, J.*, et al.* (2007). Structural basis for interaction of the ribosome with the switch regions of GTP-bound elongation factors. Molecular cell *25*, 751-764.

Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Current opinion in cell biology *17*, 316-325.

Cornish, P.V., Ermolenko, D.N., Noller, H.F., and Ha, T. (2008). Spontaneous intersubunit rotation in single ribosomes. Molecular cell *30*, 578-588.

Crick, F.H. (1968). The origin of the genetic code. Journal of molecular biology *38*, 367-379.

de Rocquigny, H., El Meshri, S.E., Richert, L., Didier, P., Darlix, J.L., and Mely, Y. (2014). Role of the nucleocapsid region in HIV-1 Gag assembly as investigated by quantitative fluorescence-based microscopy. Virus research.

Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M., and Yusupova, G. (2012). A new understanding of the decoding principle on the ribosome. Nature *484*, 256-259.

Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M., and Yusupova, G. (2013). New structural insights into the decoding mechanism: translation infidelity via a G.U pair with Watson-Crick geometry. FEBS letters *587*, 1848-1857.

des Georges, A., Hashem, Y., Unbehaun, A., Grassucci, R.A., Taylor, D., Hellen, C.U., Pestova, T.V., and Frank, J. (2014). Structure of the mammalian ribosomal pre-termination complex associated with eRF1.eRF3.GDPNP. Nucleic acids research *42*, 3409-3418.

Deutscher, M.P., and Ni, R.C. (1982). Purification of a low molecular weight form of rat liver arginyl-tRNA synthetase. The Journal of biological chemistry *257*, 6003-6006.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561-564.

Dorner, S., Brunelle, J.L., Sharma, D., and Green, R. (2006). The hybrid state of tRNA binding is an authentic translation elongation intermediate. Nature structural & molecular biology *13*, 234-241.

Dubochet, J., Booy, F.P., Freeman, R., Jones, A.V., and Walter, C.A. (1981). Low temperature electron microscopy. Annual review of biophysics and bioengineering *10*, 133-149.

Dudock, B.S., Katz, G., Taylor, E.K., and Holley, R.W. (1969). Primary structure of wheat germ phenylalanine transfer RNA. Proceedings of the National Academy of Sciences of the United States of America *62*, 941-945.

Ebihara, K., and Nakamura, Y. (1999). C-terminal interaction of translational release factors eRF1 and eRF3 of fission yeast: G-domain uncoupled binding and the role of conserved amino acids. RNA *5*, 739-750.

Eiler, D., Lin, J., Simonetti, A., Klaholz, B.P., and Steitz, T.A. (2013). Initiation factor 2 crystal structure reveals a different domain organization from eukaryotic initiation factor 5B and mechanism among translational GTPases. Proceedings of the National Academy of Sciences of the United States of America *110*, 15662-15667.

Engeland, C.E., Oberwinkler, H., Schumann, M., Krause, E., Muller, G.A., and Krausslich, H.G. (2011). The cellular protein lyric interacts with HIV-1 Gag. Journal of virology *85*, 13322-13332.

Evans, P. (2006). Scaling and assessment of data quality. Acta crystallographica Section D, Biological crystallography *62*, 72-82.

Fernandez, I.S., Bai, X.C., Murshudov, G., Scheres, S.H., and Ramakrishnan, V. (2014). Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. Cell *157*, 823-831.

Finch, A.J., Hilcenko, C., Basse, N., Drynan, L.F., Goyenechea, B., Menne, T.F., Gonzalez Fernandez, A., Simpson, P., D'Santos, C.S., Arends, M.J.*, et al.* (2011). Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. Genes & development *25*, 917-929.

Franckenberg, S., Becker, T., and Beckmann, R. (2012). Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. Current opinion in structural biology *22*, 786-796.

Frank, J. (1990). Classification of macromolecular assemblies studied as 'single particles'. Quarterly reviews of biophysics *23*, 281-329.

Frank, J., and Agrawal, R.K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature *406*, 318-322.

Frank, J., Gao, H., Sengupta, J., Gao, N., and Taylor, D.J. (2007). The process of mRNA-tRNA translocation. Proceedings of the National Academy of Sciences of the United States of America *104*, 19671-19678.

Freistroffer, D.V., Kwiatkowski, M., Buckingham, R.H., and Ehrenberg, M. (2000). The accuracy of codon recognition by polypeptide release factors. Proceedings of the National Academy of Sciences of the United States of America *97*, 2046-2051.

Freistroffer, D.V., Pavlov, M.Y., MacDougall, J., Buckingham, R.H., and Ehrenberg, M. (1997). Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. The EMBO journal *16*, 4126-4133.

Frolova, L., Le Goff, X., Rasmussen, H.H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.L., Celis, J.E., Philippe, M.*, et al.* (1994). A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. Nature *372*, 701-703.

Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. RNA *2*, 334-341.

Frolova, L., Seit-Nebi, A., and Kisselev, L. (2002). Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. RNA *8*, 129-136.

Funakoshi, Y., Doi, Y., Hosoda, N., Uchida, N., Osawa, M., Shimada, I., Tsujimoto, M., Suzuki, T., Katada, T., and Hoshino, S. (2007). Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. Genes & development *21*, 3135-3148.

Gao, H., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S.M., Van Roey, P., Agrawal, R.K., Harvey, S.C., Sali, A.*, et al.* (2003). Study of the structural dynamics of the E coli 70S ribosome using real-space refinement. Cell *113*, 789-801.

Gao, H., Zhou, Z., Rawat, U., Huang, C., Bouakaz, L., Wang, C., Cheng, Z., Liu, Y., Zavialov, A., Gursky, R.*, et al.* (2007). RF3 induces ribosomal conformational changes responsible for dissociation of class I release factors. Cell *129*, 929-941.

Gao, N., Zavialov, A.V., Li, W., Sengupta, J., Valle, M., Gursky, R.P., Ehrenberg, M., and Frank, J. (2005). Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies. Molecular cell *18*, 663-674.

Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Cote, M., Rich, R.L.*, et al.* (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell *107*, 55-65.

Godet, J., Boudier, C., Humbert, N., Ivanyi-Nagy, R., Darlix, J.L., and Mely, Y. (2012). Comparative nucleic acid chaperone properties of the nucleocapsid protein NCp7 and Tat protein of HIV-1. Virus research *169*, 349-360.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic acids research *38*, W695-699.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schafer, T., Kuster, B., Tschochner, H., Tollervey, D*., et al.* (2002). 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. Molecular cell *10*, 105-115.

Halic, M., Becker, T., Pool, M.R., Spahn, C.M., Grassucci, R.A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. Nature *427*, 808-814.

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell *107*, 679-688.

Hashem, Y., des Georges, A., Fu, J., Buss, S.N., Jossinet, F., Jobe, A., Zhang, Q., Liao, H.Y., Grassucci, R.A., Bajaj, C*., et al.* (2013). High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. Nature *494*, 385-389.

Hashimoto, Y., Kumagai, N., Hosoda, N., and Hoshino, S. (2014). The processed isoform of the translation termination factor eRF3 localizes to the nucleus to interact with the ARF tumor suppressor. Biochemical and biophysical research communications *445*, 639-644.

Hauryliuk, V., Zavialov, A., Kisselev, L., and Ehrenberg, M. (2006). Class-1 release factor eRF1 promotes GTP binding by class-2 release factor eRF3. Biochimie *88*, 747-757.

Hetrick, B., Lee, K., and Joseph, S. (2009). Kinetics of stop codon recognition by release factor 1. Biochemistry *48*, 11178-11184.

Hinnebusch, A.G. (2011). Molecular mechanism of scanning and start codon selection in eukaryotes. Microbiology and molecular biology reviews : MMBR *75*, 434-467, first page of table of contents.

Holley, R.W. (1965). Structure of an alanine transfer ribonucleic acid. JAMA : the journal of the American Medical Association *194*, 868-871.

Holley, R.W., Apgar, J., Everett, G.A., Madison, J.T., Marquisee, M., Merrill, S.H., Penswick, J.R., and Zamir, A. (1965). Structure of a Ribonucleic Acid. Science *147*, 1462-1465.

Hoshino, S. (2012). Mechanism of the initiation of mRNA decay: role of eRF3 family G proteins. Wiley interdisciplinary reviews RNA *3*, 743-757.

Ito, K., Ebihara, K., and Nakamura, Y. (1998). The stretch of C-terminal acidic amino acids of translational release factor eRF1 is a primary binding site for eRF3 of fission yeast. RNA *4*, 958-972.

Ito, K., Uno, M., and Nakamura, Y. (2000). A tripeptide 'anticodon' deciphers stop codons in messenger RNA. Nature *403*, 680-684.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nature reviews Molecular cell biology *11*, 113-127.

Jacoli, G.G., Ronald, W.P., and Lavkulich, L. (1973). Inhibition of ribonuclease activity by bentonite. Canadian journal of biochemistry *51*, 1558-1565.

Jager, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., McGovern, K.E., Clarke, S.C., Shales, M., Mercenne, G., Pache, L., Li, K*., et al.* (2012). Global landscape of HIV-human protein complexes. Nature *481*, 365-370.

Jakobsen, C.G., Segaard, T.M., Jean-Jean, O., Frolova, L., and Justesen, J. (2001). [Identification of a novel termination release factor eRF3b expressing the eRF3 activity in vitro and in vivo]. Molekuliarnaia biologiia *35*, 672-681.

Jan, E., and Sarnow, P. (2002). Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. Journal of molecular biology *324*, 889-902.

Jenner, L., Demeshkina, N., Yusupova, G., and Yusupov, M. (2010). Structural rearrangements of the ribosome at the tRNA proofreading step. Nature structural & molecular biology *17*, 1072-1078.

Jin, H., Kelley, A.C., Loakes, D., and Ramakrishnan, V. (2010). Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. Proceedings of the National Academy of Sciences of the United States of America *107*, 8593-8598.

Jin, H., Kelley, A.C., and Ramakrishnan, V. (2011). Crystal structure of the hybrid state of ribosome in complex with the guanosine triphosphatase release factor 3. Proceedings of the National Academy of Sciences of the United States of America *108*, 15798-15803.

Khatter, H., Myasnikov, A.G., Mastio, L., Billas, I.M.L., Birck, C., Stella, S., and Klaholz, B.P. (2014). Purification, characterization and crystallization of the human 80S ribosome. Nucleic acids research.

Kieft, J.S., Costantino, D.A., Filbin, M.E., Hammond, J., and Pfingsten, J.S. (2007). Structural methods for studying IRES function. Methods in enzymology *430*, 333-371.

Kim, W., Tang, Y., Okada, Y., Torrey, T.A., Chattopadhyay, S.K., Pfleiderer, M., Falkner, F.G., Dorner, F., Choi, W., Hirokawa, N.*, et al.* (1998). Binding of murine leukemia virus Gag polyproteins to KIF4, a microtubule-based motor protein. Journal of virology *72*, 6898-6901.

Kjeldgaard, M., and Nyborg, J. (1992). Refined structure of elongation factor EF-Tu from Escherichia coli. Journal of molecular biology *223*, 721-742.

Klaholz, B.P., Myasnikov, A.G., and Van Heel, M. (2004a). Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature *427*, 862-865.

Klaholz, B.P., Myasnikov, A.G., and van Heel, M. (2004b). Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature *427*, 862-865.

Klaholz, B.P., Pape, T., Zavialov, A.V., Myasnikov, A.G., Orlova, E.V., Vestergaard, B., Ehrenberg, M., and van Heel, M. (2003). Structure of the Escherichia coli ribosomal termination complex with release factor 2. Nature *421*, 90-94.

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S., and Ban, N. (2011). Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. Science *334*, 941-948.

Kong, C., Ito, K., Walsh, M.A., Wada, M., Liu, Y., Kumar, S., Barford, D., Nakamura, Y., and Song, H. (2004). Crystal structure and functional analysis of the eukaryotic class II release factor eRF3 from S. pombe. Molecular cell *14*, 233-245.

Kononenko, A.V., Mitkevich, V.A., Atkinson, G.C., Tenson, T., Dubovaya, V.I., Frolova, L.Y., Makarov, A.A., and Hauryliuk, V. (2010). GTP-dependent structural rearrangement of the eRF1:eRF3 complex and eRF3 sequence motifs essential for PABP binding. Nucleic acids research *38*, 548-558.

Kononenko, A.V., Mitkevich, V.A., Dubovaya, V.I., Kolosov, P.M., Makarov, A.A., and Kisselev, L.L. (2008). Role of the individual domains of translation termination factor eRF1 in GTP binding to eRF3. Proteins *70*, 388-393.

Korostelev, A., Asahara, H., Lancaster, L., Laurberg, M., Hirschi, A., Zhu, J., Trakhanov, S., Scott, W.G., and Noller, H.F. (2008a). Crystal structure of a translation termination complex formed with release factor RF2. Proceedings of the National Academy of Sciences of the United States of America *105*, 19684-19689.

Korostelev, A., Ermolenko, D.N., and Noller, H.F. (2008b). Structural dynamics of the ribosome. Current opinion in chemical biology *12*, 674-683.

Korostelev, A., Zhu, J., Asahara, H., and Noller, H.F. (2010). Recognition of the amber UAG stop codon by release factor RF1. The EMBO journal *29*, 2577-2585.

Korostelev, A.A. (2011). Structural aspects of translation termination on the ribosome. RNA *17*, 1409-1421.

Kos, M., and Tollervey, D. (2010). Yeast pre-rRNA processing and modification occur cotranscriptionally. Molecular cell *37*, 809-820.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell *31*, 147-157.

Kuhlbrandt, W. (1987). Three-dimensional crystals of the light-harvesting chlorophyll a/b protein complex from pea chloroplasts. Journal of molecular biology *194*, 757-762.

Kuhlbrandt, W., and Downing, K.H. (1989). Two-dimensional structure of plant light-harvesting complex at 3.7 A [corrected] resolution by electron crystallography. Journal of molecular biology *207*, 823-828.

Kushnirov, V.V., Ter-Avanesyan, M.D., Telckov, M.V., Surguchov, A.P., Smirnov, V.N., and Inge-Vechtomov, S.G. (1988). Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces cerevisiae. Gene *66*, 45-54.

Lacombe, T., Garcia-Gomez, J.J., de la Cruz, J., Roser, D., Hurt, E., Linder, P., and Kressler, D. (2009). Linear ubiquitin fusion to Rps31 and its subsequent cleavage are required for the efficient production and functional integrity of 40S ribosomal subunits. Molecular microbiology *72*, 69-84.

Lafontaine, D.L. (2010). A 'garbage can' for ribosomes: how eukaryotes degrade their ribosomes. Trends in biochemical sciences *35*, 267-277.

Lang, K., Erlacher, M., Wilson, D.N., Micura, R., and Polacek, N. (2008). The role of 23S ribosomal RNA residue A2451 in peptide bond synthesis revealed by atomic mutagenesis. Chemistry & biology *15*, 485-492.

Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., and Noller, H.F. (2008). Structural basis for translation termination on the 70S ribosome. Nature *454*, 852-857.

Lebowitz, J., Lewis, M.S., and Schuck, P. (2002). Modern analytical ultracentrifugation in protein science: a tutorial review. Protein science : a publication of the Protein Society *11*, 2067-2079.

Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., and Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic acids research *30*, 5382-5390.

Lingappa, J.R., Dooher, J.E., Newman, M.A., Kiser, P.K., and Klein, K.C. (2006). Basic residues in the nucleocapsid domain of Gag are required for interaction of HIV-1 gag with ABCE1 (HP68), a cellular protein important for HIV-1 capsid assembly. The Journal of biological chemistry *281*, 3773-3784.

Lomakin, I.B., and Steitz, T.A. (2013). The initiation of mammalian protein synthesis and mRNA scanning mechanism. Nature *500*, 307-311.

Lu, J., and Deutsch, C. (2005). Folding zones inside the ribosomal exit tunnel. Nature structural & molecular biology *12*, 1123-1129.

Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. Journal of structural biology *128*, 82-97.

Ma, B., and Nussinov, R. (2004). Release factors eRF1 and RF2: a universal mechanism controls the large conformational changes. The Journal of biological chemistry *279*, 53875-53885.

Maegley, K.A., Admiraal, S.J., and Herschlag, D. (1996). Ras-catalyzed hydrolysis of GTP: a new perspective from model studies. Proceedings of the National Academy of Sciences of the United States of America *93*, 8160-8166.

Mantsyzov, A.B., Ivanova, E.V., Birdsall, B., Alkalaeva, E.Z., Kryuchkova, P.N., Kelly, G., Frolova, L.Y., and Polshakov, V.I. (2010). NMR solution structure and function of the C-terminal domain of eukaryotic class 1 polypeptide chain release factor. The FEBS journal *277*, 2611-2627.

Marcus, A., Weeks, D.P., Leis, J.P., and Keller, E.B. (1970). Protein chain initiation by methionyl-tRNA in wheat embryo. Proceedings of the National Academy of Sciences of the United States of America *67*, 1681-1687.

Martin-Marcos, P., Cheung, Y.N., and Hinnebusch, A.G. (2011). Functional elements in initiation factors 1, 1A, and 2beta discriminate against poor AUG context and non-AUG start codons. Molecular and cellular biology *31*, 4814-4831.

Matasova, N.B., Myltseva, S.V., Zenkova, M.A., Graifer, D.M., Vladimirov, S.N., and Karpova, G.G. (1991). Isolation of ribosomal subunits containing intact rRNA from human placenta: estimation of functional activity of 80S ribosomes. Analytical biochemistry *198*, 219-223.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

McPherson, A., and Gavira, J.A. (2014). Introduction to protein crystallization. Acta crystallographica Section F, Structural biology communications *70*, 2-20.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nature structural & molecular biology *19*, 560-567.

Mitkevich, V.A., Kononenko, A.V., Petrushanko, I.Y., Yanvarev, D.V., Makarov, A.A., and Kisselev, L.L. (2006). Termination of translation in eukaryotes is mediated by the quaternary eRF1*eRF3*GTP*Mg2+ complex. The biological roles of eRF3 and prokaryotic RF3 are profoundly distinct. Nucleic acids research *34*, 3947-3954.

Moazed, D., and Noller, H.F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. Nature *342*, 142-148.

Mora, L., Heurgue-Hamard, V., Champ, S., Ehrenberg, M., Kisselev, L.L., and Buckingham, R.H. (2003). The essential role of the invariant GGQ motif in the function and stability in vivo of bacterial release factors RF1 and RF2. Molecular microbiology *47*, 267-275.

Myasnikov, A.G., Simonetti, A., Marzi, S., and Klaholz, B.P. (2009). Structure-function insights into prokaryotic and eukaryotic translation initiation. Current opinion in structural biology *19*, 300-309.

Nakamura, Y., and Ito, K. (2002). A tripeptide discriminator for stop codon recognition. FEBS letters *514*, 30-33.

Nakamura, Y., Ito, K., Matsumura, K., Kawazu, Y., and Ebihara, K. (1995). Regulation of translation termination: conserved structural motifs in bacterial and eukaryotic polypeptide release factors. Biochemistry and cell biology = Biochimie et biologie cellulaire *73*, 1113-1122.

Narla, A., and Ebert, B.L. (2010). Ribosomopathies: human disorders of ribosome dysfunction. Blood *115*, 3196-3205.

Narla, A., and Ebert, B.L. (2011). Translational medicine: ribosomopathies. Blood *118*, 4300-4301.

Ng, J.D., Gavira, J.A., and Garcia-Ruiz, J.M. (2003). Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. Journal of structural biology *142*, 218-231.

Nilsson, J., Sengupta, J., Frank, J., and Nissen, P. (2004). Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. EMBO reports *5*, 1137-1141.

Nissan, T.A., Bassler, J., Petfalski, E., Tollervey, D., and Hurt, E. (2002). 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. The EMBO journal *21*, 5539-5547.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science *289*, 920-930.

Nolan, R.D., and Arnstein, H.R. (1969). The dissociation of rabbit reticulocyte ribosomes with EDTA and the location of messenger ribonucleic acid. European journal of biochemistry / FEBS *9*, 445-450.

Noller, H.F., Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Lancaster, L., Dallas, A., Fredrick, K., Earnest, T.N., and Cate, J.H. (2001). Structure of the ribosome at 5.5 A resolution and its interactions with functional ligands. Cold Spring Harbor symposia on quantitative biology *66*, 57-66.

Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science *292*, 897-902.

Ogle, J.M., Murphy, F.V., Tarry, M.J., and Ramakrishnan, V. (2002). Selection of tRNA by the ribosome requires a transition from an open to a closed form. Cell *111*, 721-732.

Orlova, E.V. (2000). Structural analysis of non-crystalline macromolecules: the ribosome. Acta crystallographica Section D, Biological crystallography *56*, 1253-1258.

Osawa, M., Hosoda, N., Nakanishi, T., Uchida, N., Kimura, T., Imai, S., Machiyama, A., Katada, T., Hoshino, S., and Shimada, I. (2012). Biological role of the two overlapping poly(A)-binding protein interacting motifs 2 (PAM2) of eukaryotic releasing factor eRF3 in mRNA decay. RNA *18*, 1957-1967.

Pace, N.R., and Marsh, T.L. (1985). RNA catalysis and the origin of life. Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life *16*, 97-116.
Pallesen, J., Hashem, Y., Korkmaz, G., Koripella, R.K., Huang, C., Ehrenberg, M., Sanyal, S., and Frank, J. (2013). Cryo-EM visualization of the ribosome in termination complex with apo-RF3 and RF1. eLife *2*, e00411.
Petry, S., Brodersen, D.E., Murphy, F.V.t., Dunham, C.M., Selmer, M., Tarry, M.J., Kelley, A.C., and Ramakrishnan, V. (2005). Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. Cell *123*, 1255-1266.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry *25*, 1605-1612.
Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Molecular cell *37*, 196-210.
Pisareva, V.P., Pisarev, A.V., Hellen, C.U., Rodnina, M.V., and Pestova, T.V. (2006). Kinetic analysis of interaction of eukaryotic release factor 3 with guanine nucleotides. The Journal of biological chemistry *281*, 40224-40235.
Popov, S., Popova, E., Inoue, M., and Gottlinger, H.G. (2008). Human immunodeficiency virus type 1 Gag engages the Bro1 domain of ALIX/AIP1 through the nucleocapsid. Journal of virology *82*, 1389-1398.

Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A., and Ban, N. (2011). Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. Science *331*, 730-736.
Rawat, U.B., Zavialov, A.V., Sengupta, J., Valle, M., Grassucci, R.A., Linde, J., Vestergaard, B., Ehrenberg, M., and Frank, J. (2003). A cryo-electron microscopic study of ribosome-bound termination factor RF2. Nature *421*, 87-90.
Rodnina, M.V., Beringer, M., and Wintermeyer, W. (2006). Mechanism of peptide bond formation on the ribosome. Quarterly reviews of biophysics *39*, 203-225.
Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. Journal of molecular biology *333*, 721-745.

Santos, N., Zhu, J., Donohue, J.P., Korostelev, A.A., and Noller, H.F. (2013). Crystal structure of the 70S ribosome bound with the Q253P mutant form of release factor RF2. Structure *21*, 1258-1263.
Saxton, W.O., and Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial cell envelope protein. Journal of microscopy *127*, 127-138.
Scarlett, D.J., McCaughan, K.K., Wilson, D.N., and Tate, W.P. (2003). Mapping functionally important motifs SPF and GGQ of the decoding release factor RF2 to the Escherichia coli ribosome by hydroxyl radical footprinting. Implications for macromolecular mimicry and structural changes in RF2. The Journal of biological chemistry *278*, 15095-15104.
Schafer, T., Maco, B., Petfalski, E., Tollervey, D., Bottcher, B., Aebi, U., and Hurt, E. (2006). Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. Nature *441*, 651-655.
Scheer, U., Xia, B., Merkert, H., and Weisenberger, D. (1997). Looking at Christmas trees in the nucleolus. Chromosoma *105*, 470-480.
Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. Journal of structural biology *180*, 519-530.

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., *et al.* (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. Cell *102*, 615-623.

Schmeing, T.M., Huang, K.S., Kitchen, D.E., Strobel, S.A., and Steitz, T.A. (2005a). Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. Molecular cell *20*, 437-448.

Schmeing, T.M., Huang, K.S., Strobel, S.A., and Steitz, T.A. (2005b). An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. Nature *438*, 520-524.

Schmeing, T.M., Voorhees, R.M., Kelley, A.C., Gao, Y.G., Murphy, F.V.t., Weir, J.R., and Ramakrishnan, V. (2009). The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science *326*, 688-694.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. (2005). Structures of the bacterial ribosome at 3.5 A resolution. Science *310*, 827-834.

Seidelt, B., Innis, C.A., Wilson, D.N., Gartmann, M., Armache, J.P., Villa, E., Trabuco, L.G., Becker, T., Mielke, T., Schulten, K., *et al.* (2009). Structural insight into nascent polypeptide chain-mediated translational stalling. Science *326*, 1412-1415.

Selmer, M., Dunham, C.M., Murphy, F.V.t., Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R., and Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. Science *313*, 1935-1942.

Shaw, J.J., and Green, R. (2007). Two distinct components of release factor function uncovered by nucleophile partitioning analysis. Molecular cell *28*, 458-467.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proceedings of the National Academy of Sciences of the United States of America *108*, E1392-1398.

Simonetti, A., Marzi, S., Myasnikov, A.G., Fabbretti, A., Yusupov, M., Gualerzi, C.O., and Klaholz, B.P. (2008). Structure of the 30S translation initiation complex. Nature *455*, 416-420.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Song, H., Mugnier, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A., and Barford, D. (2000). The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. Cell *100*, 311-321.

Song, H., Parsons, M.R., Rowsell, S., Leonard, G., and Phillips, S.E. (1999). Crystal structure of intact elongation factor EF-Tu from Escherichia coli in GDP conformation at 2.05 A resolution. Journal of molecular biology *285*, 1245-1256.

Spahn, C.M., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., and Frank, J. (2001). Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions. Cell *107*, 373-386.

Spahn, C.M., Gomez-Lorenzo, M.G., Grassucci, R.A., Jorgensen, R., Andersen, G.R., Beckmann, R., Penczek, P.A., Ballesta, J.P., and Frank, J. (2004a). Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. The EMBO journal *23*, 1008-1019.

Spahn, C.M., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P., and Frank, J. (2004b). Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. Cell *118*, 465-475.

Spirin, A.S. (2009). The ribosome as a conveying thermal ratchet machine. The Journal of biological chemistry *284*, 21103-21119.

Spirin, A.S., Baranov, V.I., Polubesov, G.S., Serdyuk, I.N., and May, R.P. (1987). Translocation makes the ribosome less compact. Journal of molecular biology *194*, 119-126.

Stansfield, I., Jones, K.M., Kushnirov, V.V., Dagkesamanskaya, A.R., Poznyakovski, A.I., Paushkin, S.V., Nierras, C.R., Cox, B.S., Ter-Avanesyan, M.D., and Tuite, M.F. (1995). The products of the SUP45 (eRF1) and SUP35 genes interact to mediate translation termination in Saccharomyces cerevisiae. The EMBO journal *14*, 4365-4373.
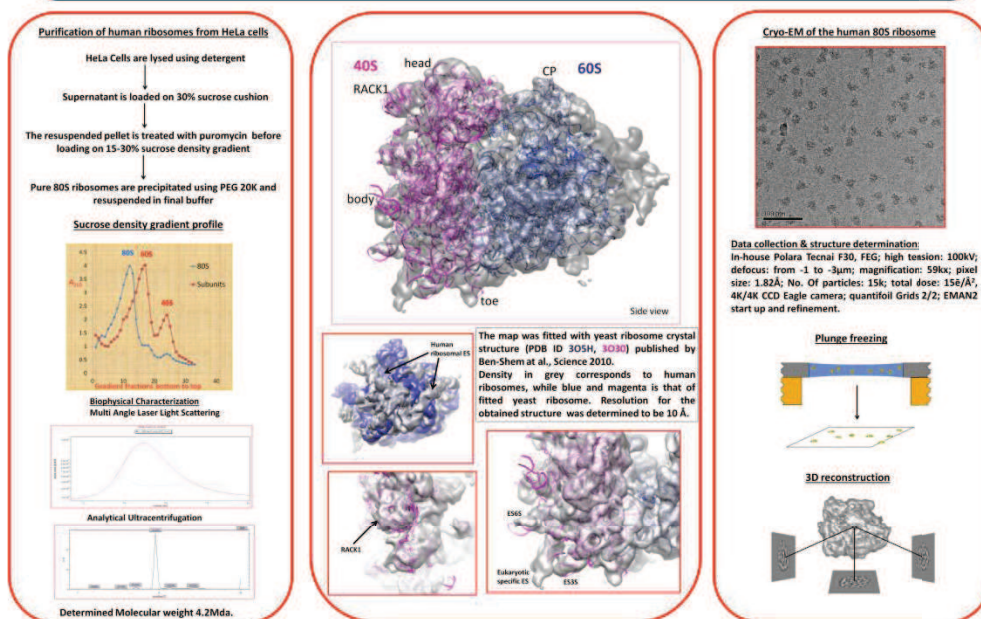
Stark, H., Orlova, E.V., Rinke-Appel, J., Junke, N., Mueller, F., Rodnina, M., Wintermeyer, W., Brimacombe, R., and van Heel, M. (1997). Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. Cell *88*, 19-28.

Stark, H., Rodnina, M.V., Wieden, H.J., van Heel, M., and Wintermeyer, W. (2000). Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. Cell *100*, 301-309.

Steitz, T.A., and Moore, P.B. (2003). RNA, the first macromolecular catalyst: the ribosome is a ribozyme. Trends in biochemical sciences *28*, 411-418.

Sweeney, R., Chen, L., and Yao, M.C. (1994). An rRNA variable region has an evolutionarily conserved essential role despite sequence divergence. Molecular and cellular biology *14*, 4203-4215.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. Journal of structural biology *157*, 38-46.

Taylor, D., Unbehaun, A., Li, W., Das, S., Lei, J., Liao, H.Y., Grassucci, R.A., Pestova, T.V., and Frank, J. (2012). Cryo-EM structure of the mammalian eukaryotic release factor eRF1-eRF3-associated termination complex. Proceedings of the National Academy of Sciences of the United States of America *109*, 18413-18418.

Tenson, T., and Ehrenberg, M. (2002). Regulatory nascent peptides in the ribosomal tunnel. Cell *108*, 591-594.

Thaller, C., Eichele, G., Weaver, L.H., Wilson, E., Karlsson, R., and Jansonius, J.N. (1985). Diffraction methods for biological macromolecules. Seed enlargement and repeated seeding. Methods in enzymology *114*, 132-135.

Thomas, J.A., and Gorelick, R.J. (2008). Nucleocapsid protein function in early infection processes. Virus research *134*, 39-63.

Trobro, S., and Aqvist, J. (2009). Mechanism of the translation termination reaction on the ribosome. Biochemistry *48*, 11296-11303.

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Molecular cell *46*, 518-529.

Tyulkina, L.G., and Mankin, A.S. (1984). Inhibition of ribonuclease contamination in preparations of T4 RNA ligase, polynucleotide kinase, and bacterial alkaline phosphatase with bentonite. Analytical biochemistry *138*, 285-290.

Uchida, N., Hoshino, S., and Katada, T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. The Journal of biological chemistry *279*, 1383-1391.

Ulbrich, C., Diepholz, M., Bassler, J., Kressler, D., Pertschy, B., Galani, K., Bottcher, B., and Hurt, E. (2009). Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. Cell *138*, 911-922.

Unwin, N., and Henderson, R. (1984). The structure of proteins in biological membranes. Scientific American *250*, 78-94.

Vagenende, V., Yap, M.G., and Trout, B.L. (2009). Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. Biochemistry *48*, 11084-11096.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M.*, et al.* (2000). Single-particle electron cryo-microscopy: towards atomic resolution. Quarterly reviews of biophysics *33*, 307-369.

van Heel, M., and Schatz, M. (2005). Fourier shell correlation threshold criteria. Journal of structural biology *151*, 250-262.

Vestergaard, B., Sanyal, S., Roessle, M., Mora, L., Buckingham, R.H., Kastrup, J.S., Gajhede, M., Svergun, D.I., and Ehrenberg, M. (2005). The SAXS solution structure of RF1 differs from its crystal structure and is similar to its ribosome bound cryo-EM structure. Molecular cell *20*, 929-938.

Vestergaard, B., Van, L.B., Andersen, G.R., Nyborg, J., Buckingham, R.H., and Kjeldgaard, M. (2001). Bacterial polypeptide release factor RF2 is structurally distinct from eukaryotic eRF1. Molecular cell *8*, 1375-1382.

Voorhees, R.M., Fernandez, I.S., Scheres, S.H., and Hegde, R.S. (2014). Structure of the Mammalian Ribosome-Sec61 Complex to 3.4 A Resolution. Cell.

Voorhees, R.M., and Ramakrishnan, V. (2013). Structural basis of the translational elongation cycle. Annual review of biochemistry *82*, 203-236.

Voorhees, R.M., Weixlbaumer, A., Loakes, D., Kelley, A.C., and Ramakrishnan, V. (2009). Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. Nature structural & molecular biology *16*, 528-533.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. Trends in biochemical sciences *24*, 437-440.

Weinger, J.S., Parnell, K.M., Dorner, S., Green, R., and Strobel, S.A. (2004). Substrate-assisted catalysis of peptide bond formation by the ribosome. Nature structural & molecular biology *11*, 1101-1106.

Weixlbaumer, A., Jin, H., Neubauer, C., Voorhees, R.M., Petry, S., Kelley, A.C., and Ramakrishnan, V. (2008). Insights into translational termination from the structure of RF2 bound to the ribosome. Science *322*, 953-956.

Wilmouth, R.C., Edman, K., Neutze, R., Wright, P.A., Clifton, I.J., Schneider, T.R., Schofield, C.J., and Hajdu, J. (2001). X-ray snapshots of serine protease catalysis reveal a tetrahedral intermediate. Nature structural biology *8*, 689-694.

Wilson, D.N., and Doudna Cate, J.H. (2012). The structure and function of the eukaryotic ribosome. Cold Spring Harbor perspectives in biology *4*.

Wilson, D.N., Harms, J.M., Nierhaus, K.H., Schlunzen, F., and Fucini, P. (2005). Species-specific antibiotic-ribosome interactions: implications for drug development. Biological chemistry *386*, 1239-1252.

Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. Nature *407*, 327-339.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.

Wohlgemuth, I., Pohl, C., and Rodnina, M.V. (2010). Optimization of speed and accuracy of decoding in translation. The EMBO journal *29*, 3701-3709.

Xiong, Y., and Steitz, T.A. (2006). A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. Current opinion in structural biology *16*, 12-17.

Yonath, A., Leonard, K.R., Weinstein, S., and Wittmann, H.G. (1987). Approaches to the determination of the three-dimensional architecture of ribosomal particles. Cold Spring Harbor symposia on quantitative biology *52*, 729-741.

Young, D.J., Edgar, C.D., Murphy, J., Fredebohm, J., Poole, E.S., and Tate, W.P. (2010). Bioinformatic, structural, and functional analyses support release factor-like MTRF1 as a protein able to decode nonstandard stop codons beginning with adenine in vertebrate mitochondria. RNA *16*, 1146-1155.

Youngman, E.M., McDonald, M.E., and Green, R. (2008). Peptide release on the ribosome: mechanism and implications for translational control. Annual review of microbiology *62*, 353-373.

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 A resolution. Science *292*, 883-896.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. (2001). The path of messenger RNA through the ribosome. Cell *106*, 233-241.

Zavialov, A.V., Mora, L., Buckingham, R.H., and Ehrenberg, M. (2002). Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Molecular cell *10*, 789-798.

Zenkova, M.A., Matasova, N.B., Myl'tseva, S.V., and Karpova, G.G. (1991). [Isolation of ribosomal subparticles from human placenta containing intact rRNA and determination of the functional activity of the 80S ribosome]. Molekuliarnaia biologiia *25*, 91-98.

Zhou, J., Lancaster, L., Trakhanov, S., and Noller, H.F. (2012). Crystal structure of release factor RF3 trapped in the GTP state on a rotated conformation of the ribosome. RNA *18*, 230-240.

Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. The EMBO journal *14*, 4065-4072.

Structure-function study of the human ribosome

Heena Khatter[1], Alexander Myasnikov[1], Stefano Stella[1], Catherine Birck[1], Isabelle Billas[1] and Bruno Klaholz[1]

[1]*IGBMC (Institute of Genetics and of Molecular and Cellular Biology), Department of Integrative Structural Biology, 67404 Illkirch, France.*

Summary

Protein synthesis in the cell is catalyzed by the ribosome and it is regulated by protein factors that bind transiently to the ribosome during the phases of translation initiation, elongation, termination and recycling. A lot of structural studies have been done on the prokaryotic ribosome and more recently also on the eukaryotic ribosome, both by crystallography (Ben-Shem et al., 2010, 2011; Rabl et al., 2010; Klinge et al., 2011) and by cryo electron microscopy of functional ribosome complexes with tRNA, mRNA, and protein factors (e.g. Chandramouli et al, 2008; Marzi et al., 2007; Spahn et al., 2004). This project involves the detailed study of the human ribosome structure and function in contrast to that of other eukaryotic organisms, with the aim of gaining insights into the specific function of the human ribosome compared to that of bacteria which are important antibiotic targets. We describe the isolation of 80S ribosomes from HeLa cells and their biophysical characterization using analytical ultracentrifugation (AUC), MALDI mass spectrometry and multi-angle laser light scattering (MALLS). Homogenous samples were frozen in the hydrated state as thin vitreous ice films (Dubochet et al., 1988) and cryo-EM images were collected on the in-house Tecnai Polara F-30 cryo electron microscope. A preliminary 3D density map has been reconstructed from these single particle images that starts showing the prominent features of the human ribosome.

Purification of human ribosomes from HeLa cells

HeLa Cells are lysed using detergent

Supernatant is loaded on 30% sucrose cushion

The resuspended pellet is treated with puromycin before loading on 15-30% sucrose density gradient

Pure 80S ribosomes are precipitated using PEG 20K and resuspended in final buffer

Sucrose density gradient profile

Biophysical Characterization
Multi Angle Laser Light Scattering

Analytical Ultracentrifugation

Determined Molecular weight 4.2Mda.

The map was fitted with yeast ribosome crystal structure (PDB ID 3O5H, 3O30) published by Ben-Shem at al., Science 2010. Density in grey corresponds to human ribosomes, while blue and magenta is that of fitted yeast ribosome. Resolution for the obtained structure was determined to be 10 Å.

Cryo-EM of the human 80S ribosome

Data collection & structure determination:
In-house Polara Tecnai F30, FEG; high teosion: 100kV; defocus: from -1 to -3µm; magnification: 59kx; pixel size: 1.82Å; No. Of particles: 15k; total dose: 15ê/Å², 4K/4K CCD Eagle camera; quantifoil Grids 2/2; EMAN2 start up and refinement.

Plunge freezing

3D reconstruction

Future Prospects

We aim at getting a high resolution structure of the human 80S and model the expansion segments in the corresponding densities. This will provide information on the functional specificity of eukaryotic ribosomes, with a prospect of developing specific antibiotics preferentially targeting the function of the prokaryotic ribosome.

Heena KHATTER

Structure-function studies of human ribosome complexes

Résumé

Les ribosomes comprennent la machinerie traductionnelle responsable de la synthèse de protéines. L'architecture et la régulation de la traduction eucaryote, en particulier chez l'homme, fut pendant longtemps un mystère pour les biologistes. Je présente ici un protocole détaillé pour purifier de manière homogène des ribosomes à partir de cellules HeLa, pouvant être utilisés pour des études à la fois biochimiques mais également structurales. En utilisant ces ribosomes, j'ai obtenu des cristaux en forme de plaque diffractant à faible résolution, pouvant être utilisés pour de futurs travaux. Une analyse par cryo-microscopie électronique de ces ribosomes a abouti à une structure à 5 Å de résolution, permettant notamment de distinguer les structures secondaires des ARNr et des protéines, permettant ainsi la construction d'un modèle. De plus, les facteurs de terminaison eRF1 et eRF3, surexprimés en bactéries, purifiés et complexés à ces ribosomes, ont permis des premières études de la terminaison de la traduction par cryo-microscopie électronique. Parallèlement, les protéines eRF1-eRF3 en complexe ont été étudiées par cristallographie aux rayons-X, montrant de nouvelles interactions jusqu'alors jamais observées. L'ensemble de ce travail fournit donc des résultats importants pour la préparation et la description de la structure du ribosome humain, pavant la voie vers l'analyse de complexes fonctionnels, à l'instar des complexes de terminaison de la traduction pour comprendre le mécanisme moléculaire inhérent à la reconnaissance du codon stop et la libération du peptide néo-synthétisé.

Mots-clés: Ribosome, synthèse protéique, 80S humain, terminaison de la traduction.

Résumé en anglais

Ribosomes comprise the translational machinery engaged in synthesizing proteins. The architecture and translation regulation of eukaryotic especially, human ribosomes, has been an enigma for a long time. Here, I present a detailed protocol for purifying homogenous ribosomes from HeLa cells which can be used for structural as well as biochemical analysis. Using these ribosomes I obtained plate-like crystals of 80S diffracting to low resolution which can be used for future work. A cryo electron microscopy analysis of these ribosomes yielded 5 Å resolution structure with secondary structures of rRNA and protein clearly visible, allowing model building. Furthermore, these ribosomes, along with the eukaryotic release factors (eRF1 and eRF3) purified by over-expression in bacteria, formed the basis for translation termination studies using cryo electron microscopy. Simultaneously, eRF1-eRF3 protein complex was explored by X-ray crystallography revealing new interactions that have not been observed before. Together, this work provides key results on the preparation and structure description of the human ribosome, paving the way for the analysis of functional complexes such as termination complexes to explore the molecular mechanism of stop codon recognition and peptide release in eukaryotic translation termination.

Keywords: ribosome, protein synthesis, human 80S, translation termination.