Università Politecnica delle Marche

Biomolecular Sciences XXXI cycle

Doctoral thesis

Characterization of archaeal protein aIF5A: a multifunctional translation factor

Alice Romagnoli

Supervisor: Prof.ssa Anna La Teana

2018

Abstract

The protein synthesis machinery includes protein factors that are highly conserved throughout evolution. Among these are EF-P in Bacteria and a/eIF5A in Archaea and Eukarya. Both, eIF5A and EF-P, are translation elongation factors which perform the essential task to rescue ribosomes from stalling during the synthesis of proteins bearing particular sequences such as polyproline stretches. Indispensable for this action is the characteristic and unique post-translation modification: hypusination in eIF5A and β-lysinylation in EF-P, which occurs in both proteins at a corresponding residue located in the N-terminal domain.

In Eukarya, a specific lysine is modified in two enzymatic reactions catalyzed by deoxyhypusine synthase (DHS), that forms an intermediate called deoxyhypusine, and by deoxyhypusine hydroxylase (DOHH) which converts the intermediate into hypusine. However, the hypusination pathway in Archaea remains obscure because, despite the presence of the hypusinated protein, homology search of archaeal genomes indicates that only the first enzyme, DHS, is present. aIF5A genes are present in all archaeal genomes sequenced to date but information concerning the protein and its function are still fragmentary. To fill this gap, we have undertaken a structural and functional characterization of aIF5A and DHS from the crenarchaeal model organism *Sulfolobus solfataricus*.

The present work shows that aIF5A in *S. solfataricus* is hypusinated, like its eukaryal counterpart. Moreover, the recombinant protein is a monomer in solution and forms a very stable complex with deoxyhypusine synthase (DHS), the first enzyme of the hypusination pathway. The enzyme forms a tetramer in solution and is able to modify its substrate *in vitro* resulting in deoxyhypusinated aIF5A, as in Eukarya. Thus, the first step towards hypusination in Sso appears to be conserved and similar to the eukaryotic one.

Concerning aIF5A function, our data confirm an evolutionary conserved role of aIF5A as a translation factor, but they also suggest the hypothesis of a multitasking protein. We provide evidence that aIF5A in fact is endowed with an RNA-binding activity as well as an RNA degrading activity. We speculate that these two conflicting properties might be regulated by the post-translational modification status of the protein (hypusinated vs non-hypusinated) and/or by its interaction with different protein partners.

Table of contents

1. Introduction

1.1 Protein synthesis in Archaea

The 'Central Dogma' of biology, proposed in 1958 by Francis Crick, states that the instructions contained in the DNA are first copied into a molecule of RNA and then converted into a functional protein product. A key step in the dogma is the process of translation, in which ribosomes translate the information, converting a nucleotide sequence into an aminoacid sequence, representing, therefore, the link between genotype and phenotype.

Protein sythesis is one of the most complex cellular processes in all domains of life for number of components and molecular interactions, and the one using up most of cell's resources. In addition, it is also one of the most conserved, in fact a translational machinery was already present in the last universal common ancestor (LUCA) of life forms.

Each of the three primary domains has developed variations of some step of translation: initiation is the main rate-limiting step and the principal target of translational regulation, moreover, together with termination and ribosome recycling, has specific features in Archaea, Bacteria and Eukarya (Londei, 2007). Elongation phase is instead essentially invariant in all cells.

In the last twenty years, many efforts have been undertaken to highlight the protein synthesis mechanism and to characterize all the different components involved in different steps, expecially for translational initiation, whose central problem is the recognition of the mRNA start codon and the setting of the correct reading frame.

Most of the differences in the mechanism of translational initiation between eukaryotes and prokaryotes can be attributed to the different structure of their mRNAs. Prokaryotic mRNAs are usually polycistronic, unmodified at their 5' and 3' ends and, in most cases, contain the Shine-Delgarno motifs (SD) that allow the direct interaction of the small ribosomal subunits with the mRNA molecules through a sequence, in 16S rRNA, complementary to SD motif in the mRNA (called anti-SD). On the other hand, eukaryotic mRNAs are monocistronic, modified by capping at the 5' end and by polyadenylation at 3' end, and do not contain any *cis*-acting ribosome binding motifs. In Eukarya, ribosomes bind the 5'-end of the transcript with the help of translation factors, start a scanning in 3' directions of mRNA and when the initiation codon is encountered, a codon/anti-codon interaction is established.

In Archaea informations about translational machinery are very limited. The early steps in protein synthesis are probably similar to the bacterial ones, but two distint mechanisms seem to exist (Benelli *et al*, 2003): one is based on mRNA/ribosome recognition through SD/anti-SD interaction and operates with polycistronic mRNAs, the other is used for monocistronic mRNAs that are devoid of 5'-UnTraslated Region (UTR). This type of mRNA, called leaderless, cannot contact directly the ribosome, but only when a codon-anticodon interactions, which requires the presence of tRNAi, has occurred. In some archaeal species, such as halophiles and extreme thermophiles of Crenarcheota branch, mRNA leaderless are very abuntant (72% in *Haloferax volcanii* and 69% in *Sulfolobus solfataricus*). This latter mechanism has been proposed to be the evolutionary oldest mechanism of initiation, since leaderless transcripts occur in all three domains of life, and they can be translated *in vitro* by all different types of ribosomes (Grill *et al*, 2000).

The overall size of archaeal ribosomes is similar to the bacterial one, in both cases, 30S and 50S subunits form mature 70S particles, and these contain the same rRNAs molecules: 5S, 16S and 23S. Similar resistance patterns to antibiotics targeting protein sythesis also outlines a similarity between archaeal and bacterial ribosomes (Benelli and Londei, 2009). By contrast, Archaea and Eukaryotes share a greater complexity in ribosome proteins composition. Ribosomal proteins in Bacteria are 57, in Archaea 68 and in Eukarya 78, among which 33 r-proteins are present only in Archaea and Eucaryotes, with high homology (Lecompte *et al*, 2002).

Besides, ribosomes must associate with specialized proteins called translation factors, which perform important functions during the translation process. Similar to ribosomal proteins, more translation factors are shared between Archaea and Eukaryotes. Translation initiation factors (IFs) help ribosomes to select start codon, interact with tRNAi and, in eukaryotes, select the cap on 5' end of mRNA. Bacteria only use three monomeric proteins: IF1, IF2, IF3; whereas eukaryotes have at least ten IFs, several of which are large multimeric complexs and completely unrelated in term of sequence homology with bacterial IFs. Even in this, Archaea seems to be closer to eukaryotes, despite the fact that they are prokaryotes. Six IFs have been identified to date in Archaea, but, as mentioned before, information about translation in these organisms is still very limited, and it is possible that additional IF proteins will be discovered in future. Nowadays we know that these six proteins are homologs to

eukaryotes IFs, and no factor is shared only by Bacteria and Archaea, underlining the evolutionary closeness of the Archaea and the Eukarya.

A list of IFs is presented in table 1:

Archaea	Bacteria	Eukarya
aIF1A	IF1	eIF1A
aIF2	IF2	eIF5B
aIF1/aSUI1	YCiH (only some phyla)	eIF1/SUI1
aIF5A	**EF-P**	**eIF5A**
a/eIF2 (trimer)	-	eIF2 (trimer)
aIF6	-	eIF6
-	IF3	-
-	-	eIF4F (trimer)
-	-	eIF3 (octamer)
-	-	eIF2B (pentamer)
-	-	eIF5

Table 1 Initiation factors in the primary domains (modified from Londei, 2005).

The probable structure of the full archaeal pre-initiation complex, including small 30S subunit, aIF1, aIF1A, Met-tRNAi, mRNA, and αßγ-aIF2 was proposed by Cryo-EM analysis (figure 1).

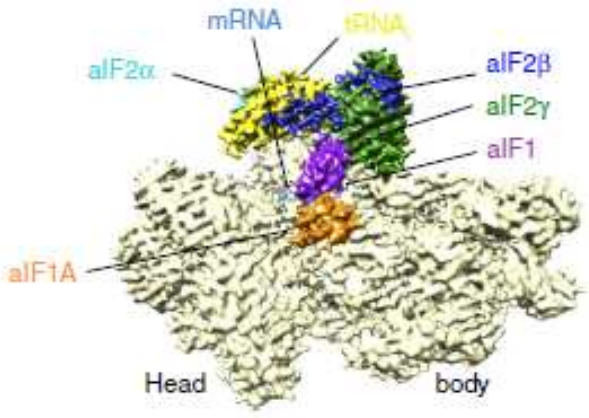


Fig.1 Cryo-EM maps of archaeal initiation complex (Coureux et al, 2016). The full archaeal initiation complex includes 30S from *Pyrococcus abyssi* (Pa-30S), shown in pale yellow, initiation factors Pa-aIF1 (in violet), Pa-aIF1A (in orange), Pa-aIF2 composed of three subunits α (in cyan), ß (in blue), γ (in green). Met-tRNAi from *E. coli* is shown in yellow and Pa-mRNA in light blue.

The trimeric aIF2, like its eukaryotic counterpart eIF2, interacts with Met-tRNAi and stimulates its binding to the 30S subunits, in a GTP-dependent reaction (Pedullà *et al*, 2005). More recently it was discovered that the γ-subunit of aIF2 is able to bind the 5' end of mRNAs, protecting them from degradation, displaying a dual function of this factors in *S. solfataricus* (Hasenöhrl *et al*, 2008). aIF1(aSUI1), homolog with eIF1, bind 30S subunits and facilitates the interaction of the tRNAi and mRNA to the ribosome (Hasenöhrl *et al*, 2006). aIF1A is a homologue of eukaryal eIF1A, but its functions still remain elusive.

The translation factor IF6 is present in Archaea and in the Eukarya, but is not found in Bacteria. Benelli and co-workers (2009) demostrated that aIF6 in thermophilic *S. solfataricus* binds specifically the large 50S subunit of ribosome preventing the formation of 70S. They also proved that aIF6 is over-expressed under stress conditions, therefore its probable function is also to modulate protein synthesis under unfavourable circumstances.

Last archaeal IF in Table 1 is the subject of this doctoral thesis: the archeal translational factor aIF5A. This factor is universal conserved and homologous between Eukarya and Archaea and with an orthologue, EF-P, in Prokaryotes.

1.2 The eukaryotic translation factor eIF5A

1.2.1 Structure and post-translational modification

The eukaryotic translation factor eIF5A is a small acidic protein of 17 kDa, highly conserved in all three primary domains, with homologues in Archaea (aIF5A) and in Bacteria (EF-P). This protein is essential in Eukarya and Archaea but not in Bacteria (Gäbel *et al*, 2013; Schnier *et al*, 1991; Park *et al,* 2010; Zhang *et al*, 2018; Balibar *et al*, 2013). It was initially identified from fractionated rabbit reticulocyte lysates as an initiation translation factor in the 1970s (Kemper *et al*, 1976; Benne *et al*, 1978).

The three- dimensional structure of human eIF5A shows that this protein is mainly composed of ß-sheets. As shown in figure 2, it consists in two different domains: a basic N-terminal domain is folded in an SH3-like barrel, found in other proteins related to translation, while the C-terminal domain harbors an OB-fold (oligonucleotide-binding fold), a five-stranded beta-barrel known to bind nucleic acids and typical of other translational machinery components, like eIF1A, eIF2α and other ribosomal proteins (Dias *et al*, 2013).

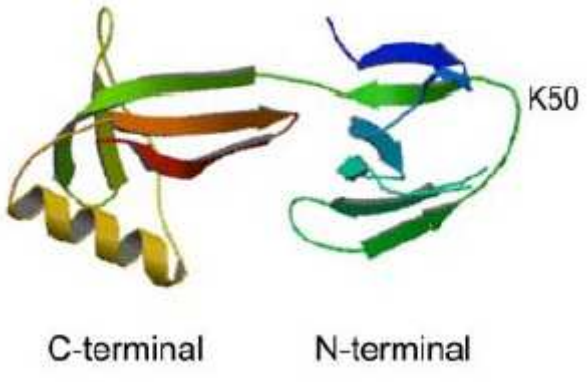


Fig. 2 Crystal structure of human eIF5A protein (PDB ID code 3cpf) (Park et al, 2010)

This eIF5A protein shows high amino acids conservation in all eukaryotes, particularly in an exposed loop linking ß3 and ß4 strands in N-terminal domain, in which a special lysine residue (K50 in human and K51 in yeast eIF5A) is post-translational modified in hypusine residue.

Fig.3 Conservation of the amino acid eIF5A sequence in Eukarya. The human sequence is shown. The degree of conservation is indicated by color coding: red, 100% identity; dark orange to yellow, conservative replacements with >80 to >50% sequence identity; white, no significant sequence identity (Wolff *et al*, 2007)

The amino acid sequence surrounding this modification site (SKTGKHGHAK) is found in more than 100 different species in eukaryotes, and is a very basic and hydrophilic portion (Park *et al*, 2010).

Hypusine was discovered by Shiba *et al.* (1971) from bovine brain extract during their search for new amine components. Its name derived from its two components: hydroxyputrescine and lysine. About 10 years later it was discovered that hypusine is present exclusively in the translational factor eIF5A, being added in a post-translational modification process. Until today eIF5A is still the only known hypusinated protein. This unique modification pathway

is achieved in two enzymatic steps (Figure 4): in the first step deoxyhypusine synthase (DHS) tranfers an aminobutyl group from polyamine spermidine to one specific lysine residue (K51 in yeast) of eIF5A precursor, catalyzing the formation of an intermediate, deoxyhypusine [N^{ε}-(4-amino-butyl)-lysine] residue. DHS depends on NAD^{+} as coenzyme. This intermediate is subsequently hydroxylated by deoxyhypusine hydroxylase (DOHH) to form mature and active eIF5A with hypusine residue (Park *et al*, 2010; Wolff *et al*, 2007).

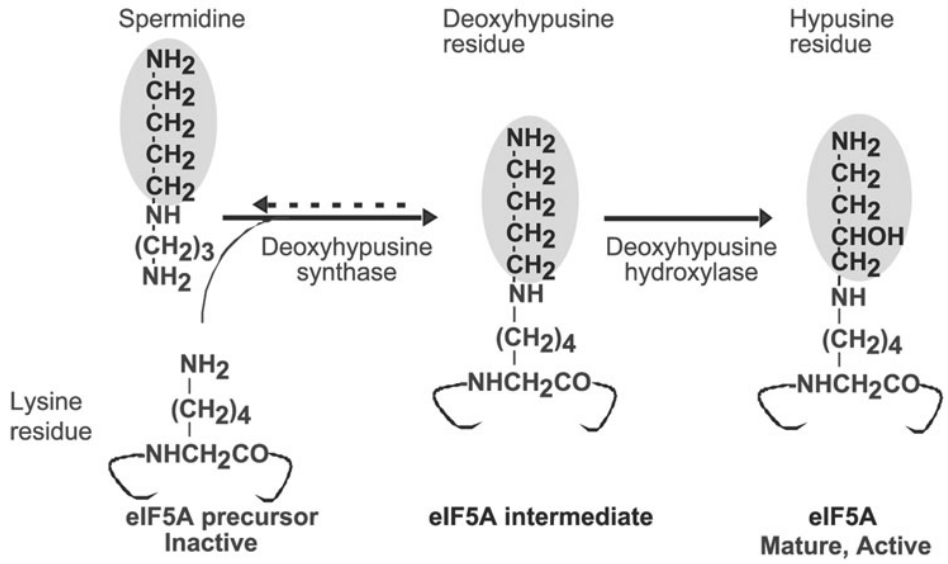


Fig.4 Hypusine biosynthesis in eIF5A (Park, 2006).

Regarding the quaternary structure of eIF5A, the protein was shown to exist as a homodimer both in solution and *in vivo.* In a study of 2009, Gentz and coworkers demostrated that formation of the eIF5A dimer in *S. cerevisiae* requires the presence of the hypusine residue as well as the presence of cellular RNA molecules (Gentz *et al*, 2009). More recent experiments always performed in yeast, showed that eIF5A forms dimers, both *in vitro* and *in vivo*, in an RNA-dependent manner, but regardless of the presence of the hypusine residue. This oligomer conformation was also confirmed by SAXS analysis which showed that the eIF5A dimer is L-shaped resembling the bacterial EF-P and superimposable with the tertiary structure of a tRNA (Dias *et al*, 2013).

eIF5A and its hypusine modification are essential for eukaryotes and besides eIF5A, also the enzymes DHS and DOHH are highly conserved in the entire eukaryotic kingdom (Wolff *et al*, 2007).

Two or more eIF5A isoforms have been identified in Eukarya (Jenkins *et al*, 2001). In yeast *S. cerevisiae* there are two eIF5A genes, *TIF51A* and *TIF51B*, regulated by oxygen with *TIF51A* mainly expressed under aerobic condition while *TIF51B* only under anaerobic conditions (Schnier *et al*, 1991). These two genes were found to perform the same function in yeast and gene inactivation of one gene or both rendered *S. cerevisiae* not viable, underling the essential nature of eIF5A. The eIF5A-depleted cells showed a phenotype characteristic of G1 arrest (Kang and Hershey, 1994). On the contrary, in higher multicellular eukaryotes like *C. elegans*, the two eIF5A genes, *IFF-2* and *IFF-1*, code for two isoforms (eIF5A1 and eIF5A2 respectively) with different functions and tissue localization. eIF5A1 is required in soma for somatic tissue growth and organization, while eIF5A2 is required in germline and gametogenesis (Hanazawa *et al*, 2004).

The essential role of hypusine and the ezymes involved in this modification, were also studied in yeast, through the costruction of eIF5A-1 mutant (K51R), in which the specific lysine 51 was substituted with arginine. DHS enzyme, is unable to modify this eIF5A mutant in *S. cerevisiae* strains, making this yeast strain unfit to grow (Schnier *et al*, 1991).

DHS structure was characterized in human through crystal structure of recombinant enzyme in complex with NAD^+, available since 1998 (Liao *et al*, 1998). Human DHS is a homotetramer composed of four identical subunits of 40 kDa and its active sites are located at the interface between dimers, providing the binding site for NAD^+. There is a single DHS gene in yeast and in most eukaryotes. The role of deoxyhypusine synthase was demostrated by gene distruption in yeast and with knock-out studies in mice. In both cases it was established the essential requirement for DHS, for growth and survival of eukaryotes, from yeast to mammals.

Like eIF5A and DHS, also DOHH is highly conserved in the eukaryotic kingdom. DOHH exists as a product of a single gene which in human is a 32 kDa protein. Sequence alignment of DOHH proteins reveals that this enzyme belongs to a HEAT-repeat-containing proteins family. It is composed of a symmetrical super helical structure with 8 helical hairpins and requires oxygen and iron for his activity, harboring in his active center a nonheme diirion site. Recently the crystal structure of human DOHH has also been published (Han *et al*.,

2015). Curiously, DOHH activity is only essential in higher eukaryotes, such as mammals, but not in yeast. In mammalian cells metal-chelating inhibitors of DOHH cause G1 cycle arrest and growth inhibition (Hanauske-Abel *et al*, 1994). Moreover, inactivation of DOHH is also lethal in *C. elegans* and *D. melanogaster* (Spradling *et al*, 1999; Maeda *et al*, 2001; Patel *et al,* 2009). In contrast, in yeast *S. cerevisiae* the DOHH gene is not essential, even though native eIF5A exists as hypusine-modified form. This may explain the recent discovery in protozoan parasite *Trichomonas vaginalis*, in which *dohh* gene is lacking but eIF5A protein is completely hyspusinated. Researchers demonstrate in this study that TvDHS is able to perform both DHS and DOHH reactions, because in its tetrameric structure that resemble the human DHS, the enzyme has also HEAT-motifs typical of DOHH enzyme, allowing to perform also hydroxylase reactions (Quintas-Granados *et al*, 2015).

In addition to the hypusine post-translational modification, it is known that yeast eIF5A is also phosphorylated on the Ser2 residue of the protein (Klier *et al*, 1993), but mutation studies reveal that phosphorylation of eIF5A does not appear to be indispensable for protein's function, unlike its hypusine modification.

1.2.2 Function of eIF5A in translation

eIF5A was originally purified from ribosomes of reticulocyte lysate in 1978, suggesting immediately an involvement in protein translation. At the beginning it was hypothesized that eIF5A stimulate the methionyl-puromycin synthesis, indicative of a role in the formation of the first peptide bond (Benne *et al*, 1978). In the years, biochemical and molecular genetic studies of yeast eIF5A provided new information about the function of this translation factor. Examining the impact of eIF5A depletion in yeast, a pronounced defect in total protein synthesis was observed, unlike the moderate inhibition previously reported (Kang and Hershey, 1994; Saini *et al,* 2009). Moreover, in other studies it was demonstrated that eIF5A fisically interacts with structural components of translation machinery. Proteins co-purifying with recombinant tagged-eIF5A were identified in *S. cerevisiae* and among these a 60S ribosomal protein P0, 40S ribosomal protein S5 and translation elongation factor 2 (eEF2) were found. These interactions appeared to be hypusine-dependent, since mutation of lysine 51 to arginine in the site that undergoes hypusination, decreased intensity of the interaction just described, especially for eEF2, as loss of interplay between eIF5A/eEIF2 was observed

(Zanelli *et al*, 2006). Recently it was proved that these two translation factors display a negative cooperativity on binding to the ribosome (Rossi *et al*, 2016).

Interaction of eIF5A with ribosomes was also confirmed by the work of Jao and Chen (2006), in which they provided a list of eIF5A-tagged interacting partners. 14 out of 19 proteins are ribosomal proteins from either 40S or 60S, the other proteins include translation factor 1A (eIF1A) and two chaperonine. They suggested that eIF5A binds 80S ribosome in active translation and this interaction requires RNA and hypusine modification. Using endogenous yeast eIF5A, Jao *et al* (2006) reached the same conclusions.

A further evidence for the involvement of eIF5A in translation elongation was obtained by polysome profile analysis. This technique is used mainly to assess translation function in yeast, with cycloheximide treatment that freezes ongoing translation elongation and extracts are then separated on sucrose gradients. Analyses of polysome profile of temperature-sensitive eIF5A mutants showed an increase of polysome/monosome ratio at the restrictive temperature, that could be explained with a block of translation elongation (Zanelli *et al,* 2006), the same pattern expected for polysome profile of wild type cells after sordarin treatment, an elongation factor 2 inhibitor (Saini *et al*, 2009), supporting once again a role of eIF5A in elongation.

An important scientific milestone for the function of eIF5A in translation was reached in 2013 when Gutierrez and coworkers used *in vivo* and *in vitro* assays to unveil that eIF5A plays a crucial role in translation of specific proteins containing consecutive polyproline residue (Gutierrez *et al,* 2013), consistent with the reports of the ortholog EF-P. A set of dual-luciferase reporter constructs in which 5' *Renilla* luciferase and 3' firefly luciferase open reading frames (ORFs) are separated by sequences encoding ten consecutive codons for each of twenty amino acids, were introduced into wild type and eIF5A temperature-sensitive eIF5A mutant strains of yeast. Monitoring of all 20 luciferase reporter constructs revealed that only those containing proline codons were specifically impaired in eIF5A mutant. *In vitro* analysis confirmed this result, as only those peptides containing at least three consecutive proline residues showed absolute dependence on eIF5A. Ribosomes encountering these polyproline motifs stall, they are unable to proceed with translation and need to be rescued by eIF5A.

The reason for this, resides in the fact that not all peptide bonds are indeed formed with equal efficiency, as certain amino acids are poor donors or acceptors. Proline is an imino acid with

its side-chain cyclized onto the backbone nitrogen. The constrained geometry of the cyclic proline side chain makes this amino acid both a poor acceptor as well as a poor donor in the peptydil transferase reaction, therefore consecutive proline stretches causes the ribosome to stall.

Structural studies of eIF5A bound to the ribosome helped to understand how eIF5A could contribute to rescuing ribosomes stalled during translation. A model was proposed in 2016 by Schmidt with cryo-electron microscopy (Schmidt *et al,* 2016) and subsequently by Melnikov with crystal structure of eIF5A bound to 80S *S. cerevisiae* ribosome (Melnikov *et al,* 2016), in which eIF5A is located between the P and the E sites, interacting with tRNAs, rRNA and ribosomal proteins. As shown in figure 5B, domain I of eIF5A contacts 25S rRNA nucleotides located in helices H74 and H93, while domain II contacts ribosomal protein L1 and L42 and charged hypusine interacts A76 of the CCA-end of peptidyl-tRNA in PTC of the ribosome, stabilizing and restricting the position of the polyprolines in the P-site. Thus, eIF5A facilitates the transfer of the nascent chain from P to A-site.

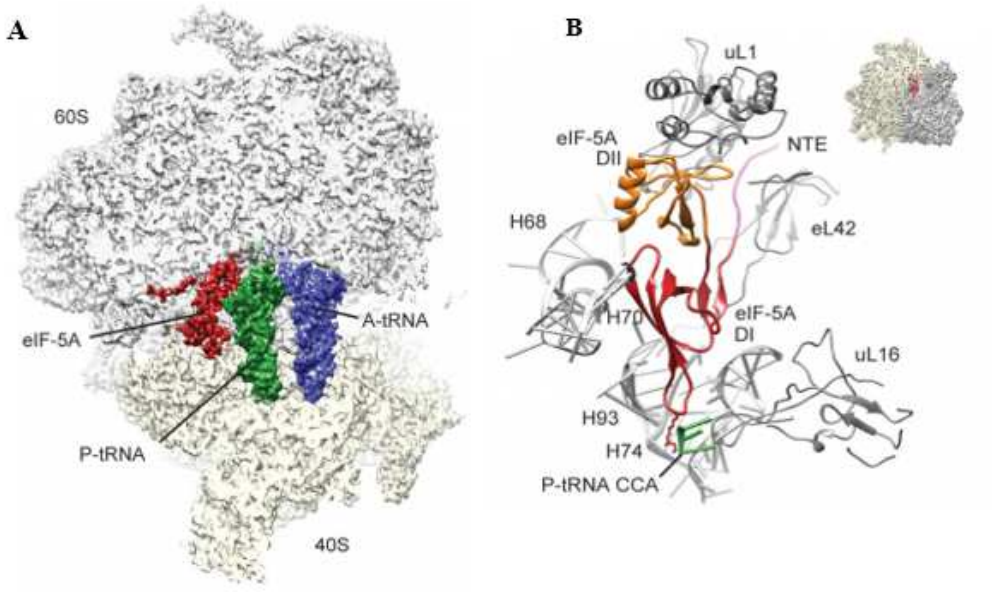


Fig.5 Cryo-EM structure of eIF5A bound to the yeast 80S ribosome (A) Cryo-EM map of eIF5A-80S complex, in which it is shown the eIF5A binding site (B) Molecular model for the interaction of domain I and II to the ribosome component. (Schmidt *et al*, 2016).

Using a proteomic approach, the frequency of polyproline motifs in all three primary domains of life was determined and it was found that the frequency of PPP or PPG containing proteins increases with organism complexity, in *E. coli* the frequency is 5.8%, in *S. cerevisiae* is 9,7% while in *H. sapiens* is 33,5%. This could explain why eIF5A is essential in all eukaryotes contrarly to EF-P of bacteria (Mandal *et al*, 2014), and the number of polyproline triplets increases considerably during the evolution (Figure 6).

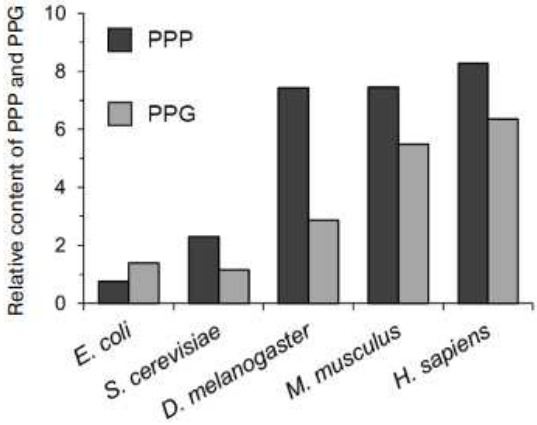


Fig.6 The frequency of PPP and PPG motifs in different organisms. The histogram shows the frequency of Pro-pro-pro and pro-pro-gly per 10^4 amino acid residues (Turpaev, 2018).

However, this view has been slightly changed in a recent study in which ribosome profiling of eIF5A depletion strain in yeast showed that many other tri-peptide motifs cause ribosome stalling. Interestingly, many of these motifs contain not only polyproline, like previously reported, but a combination of aspartic acid, glycine and other amino acids and few of these contain even no proline at all. Moreover, eIF5A depletion impacts in a defect of global elongation rate, while EF-P loss does not (Schuller *et al*, 2017). These findings led to hypotize that eIF5A is involved in a global translation elongation and support a model in which eIF5A is a general elongation factor, in contrast of the ortholog EF-P, which is much more specialized in translation of proteins containing proline stretches.

In the same study Schuller *et al.* also demostrated a role of eIF5A in translation termination, greatly expanding the role of this essential translation factor. Using the same eIF5A-depletion strain, they proved, through *in vivo* ribosome profiling experiments, that the stop

codon peak is 12-fold higher compared to wild-type and ribosome occupancy in 3'-UTR is increased, consistent with a translation termination defect, while previous studies of this group showed no such effect in *E. coli* cells lacking EF-P. They also showed using an *in vitro* translation system, that hypusinated eIF5A stimulates an eRF1-mediated peptide hydrolysis of peptidyl-tRNA by 17-fold compared to wild type, and hypusine is critical for this activity, as the rate of peptidyl-tRNA relase has a minor effect (4-fold). Taken together, this data provided support to the idea that eIF5A plays a crucial role in translation termination by stimutating the activity of eRF1.

1.2.3 Other functions of eIF5A and implication in pathological cellular process

In addition to its role as a translation factor, the eukaryal eIF5A has been implicated into a variety of cellular processes, like mRNA decay (Zuk and Jacobson 1998), cell cycle progression (Hanauske-Abel *et al*, 1994), apoptosis (Caraglia *et al*, 2013), cell polarity (Chatterjee *et al*, 2006; Zanelli e Valentini 2005), retroviral and protozoan infection (Hoque *et al*, 2009, Olsen and Connor, 2017) and stress response (Gosslau *et al*, 2009).

Studies on the immunodeficiency virus (HIV) reveal that eIF5A interacts with viral mRNAs, mediating the translocation of these mRNAs from the nucleus to the cytoplasm, and these interactions are mediated by a stable binding of the hypusinated protein with specific mRNA (Liu *et al*, 1997). Using SELEX technique (systematic evolution of ligands by exponential enrichment) it was shown that mRNAs bound by eIF5A, isolated by co-immunoprecipitation with FLAG-tagged eIF5A, shared a conserved motifs UAACCA and AAAUGU (Xu and Chen, 2001). They speculated that eIF5A-RNA interaction is due to some structural elements in RNA molecules, like hairpins and internal loops, suggesting that eukaryotic translation factor 5A is also a specific hypusine-dependent RNA binding-protein. This notion finds support in the analysis of the structure of eIF5A C-terminal domain, which, as described above, has an OB-fold typical of nucleid acid binding-protein. In particular this portion of the protein is negatively charged and shares a structural similarity with the RNA chaperone CspA.

There is also evidence that eIF5A is a key protein in the pathogenicity of different diseases, such as diabetes, malaria, HIV-1 infections and human cancer (Kaiser *et al*, 2012). Several

studies in these last years own the implication of eIF5A in many tumor types as a main topic. In particular, it has been shown that the two isoforms of human eIF5A display different functions and expression patterns. eIF5A-1 protein is constitutively expressed in all tissues, whereas eIF5A-2 is not normally detectable, but is over-expressed in certain human ovarian cancer tissues and colorectal (Clement *et al*, 2003), human hepatocarcinoma cell lines and mouse embryonic livers (Lee *et al,* 2010) and this upregulation of the protein caused cellular transformation. Based on these findings eIF5A-2 is proposed to be an oncogene. In this scenario, the hypusine modification seems to have an important task. Several studies indicate that hypusination of eIF5A is required for tumor maintenance and disease progression and the inhibition of hypusine prevents tumorogenesis (Nakanishi and Cleveland, 2016). Hence research in this field is moving towards the identification of therapeutic compounds targeting eIF5A protein and the hypusination pathway (hypusine, DHS, DOHH), which could be considered as new drugs and, combined with the other therapies, used to treat cancer patients.

1.3 The bacterial translation factor EF-P

1.3.1 Structure and post-translational modification

EF-P, isolated by Glick and Ganoza in 1975, is a small protein of 21 kDa which exibits a three-domain structure (I, II, III), with an overall L shape remindful of a tRNA molecule (figure 7). The N-terminal part consists of domain I and II, forming one arm of the L, whereas the other arm is formed by domain II and III, the latter domain forming the C-terminus of EF-P. Domains II and III are similar to the OB-fold motif (oligonucleotide-binding fold) observed in *E. coli* cold shock protein, polyribonucleotide nucleotidyltransferase, transcription factor Rho, and other proteins known to bind DNA or RNA molecules. However, EF-P is an acidic protein (PI= 4.6) and most of the overall surface is negatively charged, especially domain II and III presents most of the negatively charged aminoacids (Glu-76, Glu-78, Asp-84, Glu 89, Glu-106 for domain II and Asp-134, Glu-154, Glu-166, Glu-169 for domain III are conserved).

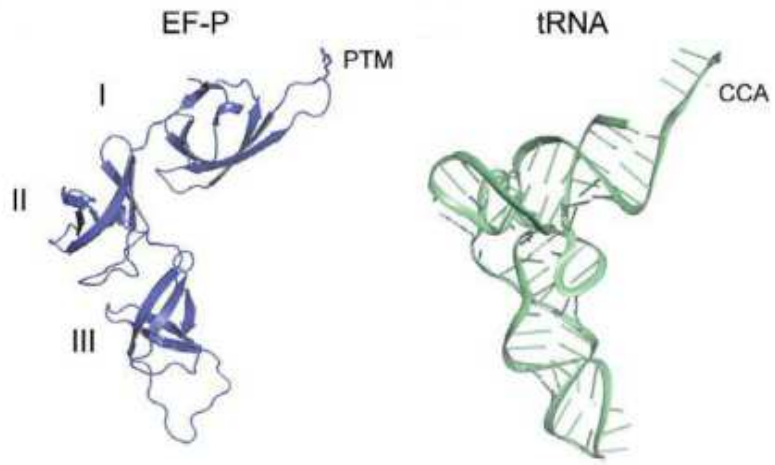


Fig.7 Comparison of crystal structure of EF-P from T. thermophilius (PDB ID code 1UEB) and tRNA (Lassak et al, 2016).

As shown in figure 8, Domain I of EF-P and the N-terminal domain of eIF5A show the same fold, as do domain II of EF-P and C-terminal domain of eIF5A. The third domain is an exclusive of EF-P and has high structural homology with domain II: this suggest that this additional domain in EF-P arose as a result of a duplication event (Lassak *et al*, 2016). Crystal structure of *T. thermophilus* of EF-P (PDB ID code 1UEB) demostrate that the angle formed by the two arms of EF-P is around 95°C and notably, the interfaces between domain I-II and II-III are composed of hydrophobic chains with high surface complementarities. This proves that the L shape of this translational factor is due to a native conformation of the protein instead of an artifact in crystal packing (Hanawa-Suetsugu *et al*, 2004). Comparison of this structure with the other X-ray results of the bacterial EF-P available on PDB, coming from *P. aeruginosa* (Choi and Choe, 2011) and *Clostridium thermocellum* (PDB ID code 1YBY), denotes a high structural flexibility (>60° rotation) of domain I respect to domains II and III, which remain relatively rigid.

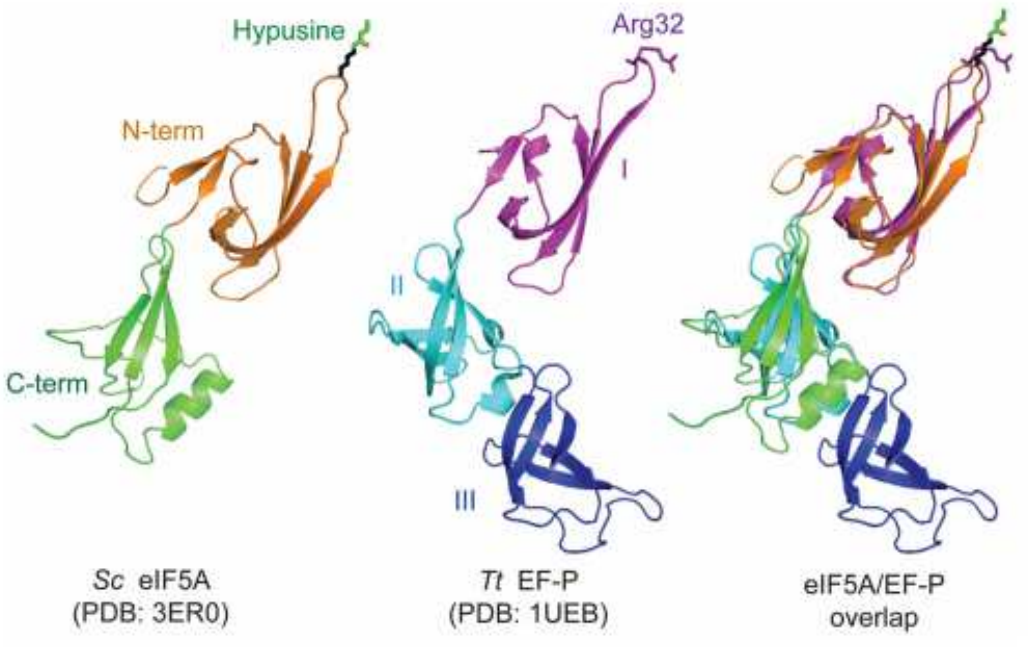


Fig.8 Comparison of eukaryal eIF5A and bacterial EF-P structure. (Dever et al, 2014).

The alignment of the bacterial EF-P sequence with eukaryal eIF5A shows that in the N-terminal domain, in the region corresponding to the eukaryal hypusinated Lys50 residue, there is the highest conservation of aminoacids, although in *E. coli* DHS and DOHH are absent (Park *et al*, 2011). In some species including *Thermus thermophilus*, lysine residue is replaced by arginine. Several experiments proved that also EF-P is subjected to a post-translational modification, although not in all bacteria. This modification, studied in *E. coli* and *Salmonella sp*, is called ß-lysylation and occurs in three steps and requires the enzymes YjeK, YjeA and YfcM (Navarre *et al*, 2010; Yanagisawa *et al*, 2010; Peil *et al*, 2012).

In a first step the enzyme YjeK (called also EpmB), a lysine-2,3-aminomutase, converts a free S-α-Lysine to R-ß-Lysine (Behshad *et al*, 2006). The second enzyme involved in this modification is a paralog of the lysyl-tRNA synthetases (YjeA, also known as EpmA) and transfers an R-ß-Lysine portion to the ϵ-amino group of a specific lysine, using ATP. Based on this hypothesis, the crystal structure of *E. coli* YjeA in complex with EF-P was solved in 2010, demostrating lysinylation of Lys 34 (Yanagisawa *et al*, 2010). Moreover, in a previous experiment, Aoki *et al*. revealed that native EF-P purified from *E. coli* had an extra mass of +144 Da at Lys34 (Aoki *et al*, 2008). Sole lysinylation is predicted to increase the mass of

EF-P of only +128 Da. This discrepancy was resolved when Peil *et al* (2012) demostrated that another step of modification occurrs, performed by YfcM (EpmC). They proved by MS-analysis that native EF-P from cells lacking YfcM showed a decrease of 16 Da in mass, indicating that this third enzyme complete the modification patway of EF-P in *E. coli* (Peil *et al*, 2012). YfcM recognizes EF-P only in its modified form and hydroxylates the lysyl-lysine residue.

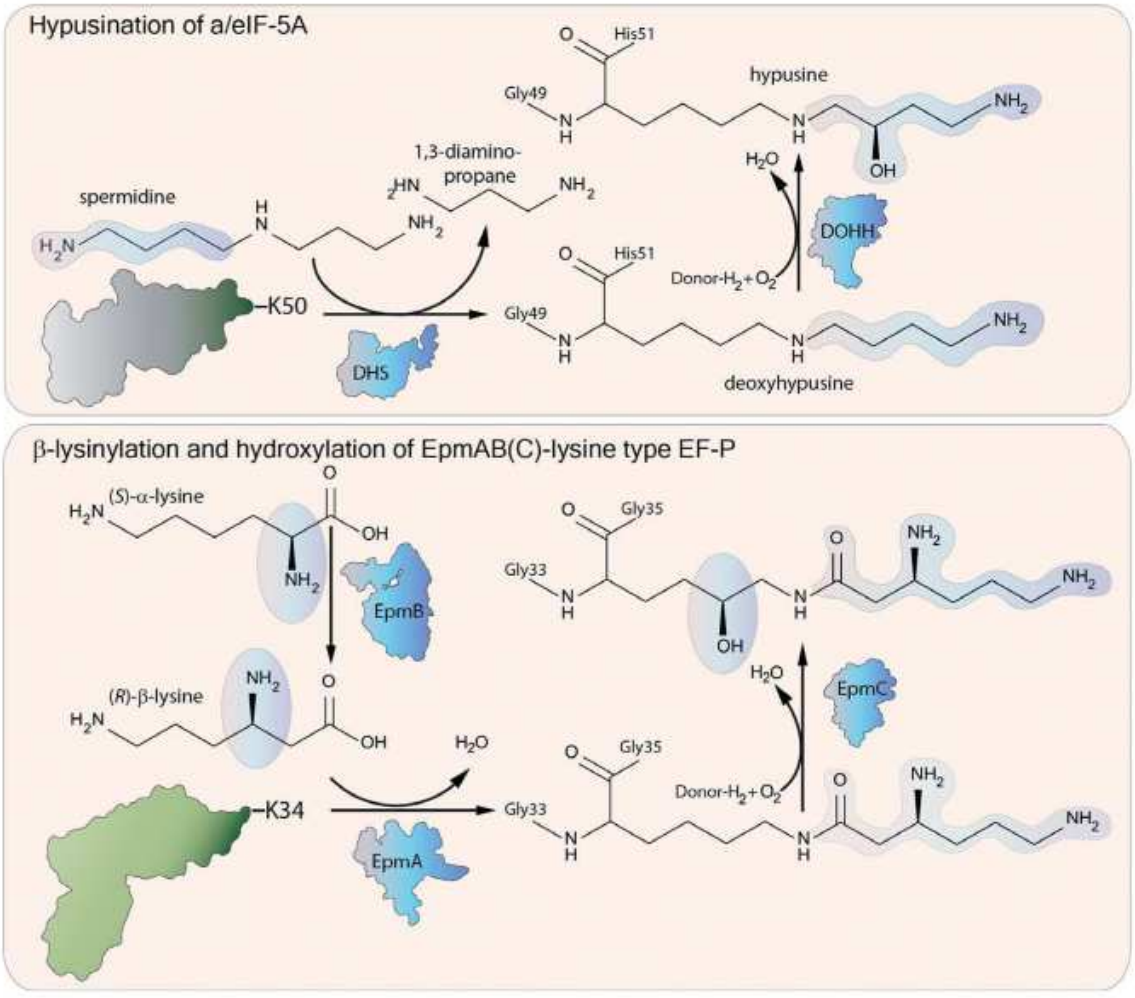


Fig.9 Diversity of EF-P/eIF5A modification systems (Lassak et al, 2016)

Although hydroxylation is the last modification step also in eukaryotes, there is no evolutionary correlation between hydroxylases involved in hypusination of eIF5A and lysinylation of EF-P.

This modification system has been identified only in about one quarter of all bacterial genomes (EpmA and EpmB are found in about 26% of bacteria while EpmC is restricted almost excusively to γ-proteobacterial genomes), but ß-lysinylation is necessary for EF-P function in species that contain this modification, such as *E. coli* and *Salmonella sp*. Based on the importance of modification in this species, Lassak and coworkers speculated that lysinylation is not the only modified mechanism present in Bacteria (Lassak *et al*, 2015). As mentioned above, some species are characterized by a conserved arginine residue in EF-P (Arg32) at the position equivalent to Lys34 found in *E. coli.*

Using γ-proteobacterium *Shewanella oneidensis* as a model organism they described rhamnosyl modification pathway, in which 2'-deoxy-thymidine-ß-L-rhamnose is attached on the arginine 34 of EF-P by a glycosyltransferase (EarP). Also, this modification produced a mass shift of +146 Da in MS-analysis in EF-P -Arg32 modification pathway, consistent with a deoxyhexose formation. Members that assess EF-P activation via EarP also comprise notable pathogenic bacteria as *Neisseria sp., Bordetella pertussis* and *Pseudomonas aeruginosa,* and deletion of *efp* or *earP* genes lead to a lack of pathogenicity in these organisms. In *P. aeruginosa* mutant strains are unable to kill human cells. (Lassak *et al,* 2015).

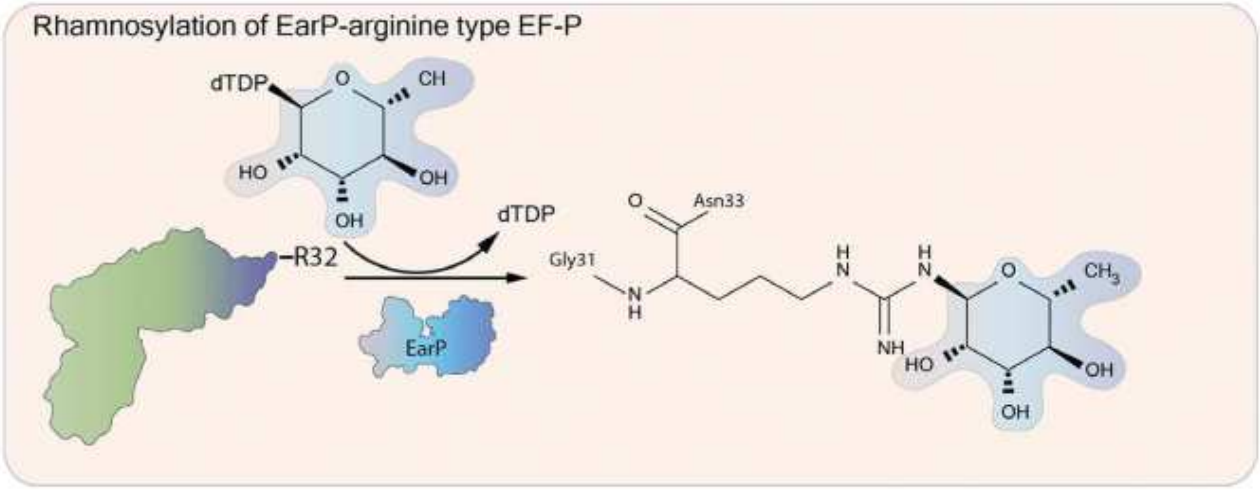


Fig.10 Rhamnosylation of EF-P in Bacteria (Lassak et al, 2016)

Anyway, it is important to mention that both modification in EF-P and the enzymes involved in these pathways were identified in only about 35% of the availabe bacterial genomes. The hypothesis is that there are probably alternative bacterial strategies of modification not identified yet.

1.3.2 Function of EF-P

EF-P is coded by only one gene, *efp*, and is universally conserved in Bacteria. In contrast to what was reported for eIF5A, EF-P is not an essential protein. It was demonstrated that EF-P is dispensable in *Escherichia coli* and *Pseudomonas aeruginosa* (Balibar *et al*, 2013). The EF-P protein was initially identified as an elongation factor, as was predominantly isolated from post-ribosomal supernatant (90%). According to this model EF-P stimulates peptide bond formation during translation elongation (Glick *et al*, 1975). Afterwards, some studies speculated on the hypotesis that EF-P enhances peptidyl-transferase activity in specific aminoacyl-tRNAs, such as those charged or with relative wide size of amino acids side chain (Glick *et al*, 1979; Ganoza and Aoki, 2000). The next important scientific breakthrough for EF-P was in 2009 when a crystal structure of *Thermus thermophilus* EF-P bound to 70S ribosome (Fig.6) was described (Blaha *et al*, 2009). Despite its resemblance to a tRNA, EF-P bound to the 70S ribosome between the P and E sites, adjacent to the P-site tRNA and interacts with both ribosomal subunits in the interface between 30S and 50S. Domain I and II of EF-P, which are superimposable to N-terminal and C-terminal of a/eIF5A, bind next to the acceptor stem and D-loop of tRNA in P site; EF-P domain III, which is absent in a/eIF5A, interacts with fMet-tRNA near its anticodon stem-loop, close to 30S subunit. EF-P domain II has also connection with L1 ribosomal protein, that is structurally and functionally conserved among all organisms, involved in tRNA translocation and release. In domain I a loop contacts the CCA-end of tRNA, maybe orienting the amino acyl arm optimally for peptide bond formation.

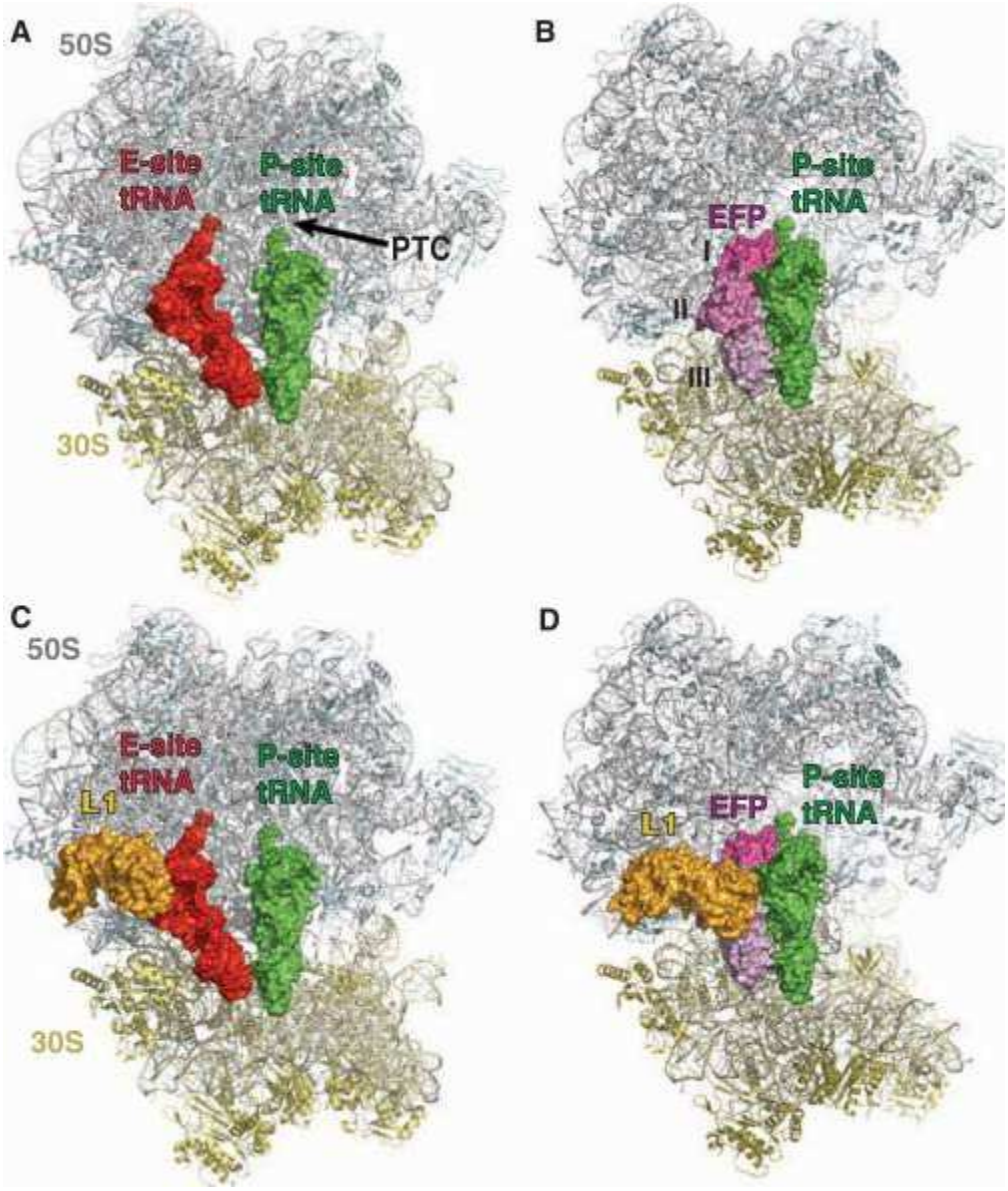


Fig.11 Crystal structure of EF-P bound to the ribosome. (A) E-site (shown in red) and P- site (shown in green) tRNAs bound to 70S. (B) EF-P (shown in shades of magenta to indicate different domain of the protein (I, II, III) and P-site tRNA binding in the 70S ribosome. (C) Same as (A) but with ribosomal protein L1 (shown in gold). (D) L1 movement, due to the presence of EF-P. (Blaha *et al*, 2009)

Next important insights for the function of EF-P were provided by studies of the eukaryotic eIF5A, that was demostrated in 2009 to have a role during translation elongation.

Mutant of EF-P, YjeA, YjeK cause the accumulation of polysomes (Bullwinkle *et al*, 2013) and these bacteria exibit pleiotropic phenotypes, correlated with specific cellular process, such as virulence, cell mobility, stress conditions resistance. These results were consistence with a function of EF-P in translation elongation, in particular it could be involved in

traslation of specific mRNAs, as described for eIF5A. In fact, recent studies in ribosome profiling of *E. coli* demostrated that EF-P is important in peptidyl bond formation between two proline motifs. In Bacteria, lacking EF-P or the enzyme for lysinylation, ribosomes stop at these motifs and are unable to procede with the translation. Levels of pausing at Pro-Pro varied depending on precedent and subsequent codon: PPW, PPN, PPD, PPP, GPP, DPP, APP induced very strong pauses whereas PPI, PPL, PPM, PPY, PPF exhibited no pausing. As previously described for the eukaryal counterpart, this phenomenon is due to the geometry of the cyclic proline side chain and EF-P function is to stabilize the CCA-end of the peptidyl-tRNA. More recently, the same poly-proline specific function for eIF5A has been demostrated in yeast by Gutierrez and coworkers (2013), suggesting that the hypusine residue directly interact the CCA-end of tRNA in PTC. The same mechanism can be proposed for EF-P in that 30% of bacteric species harboring the Lys/Arg 34. However, the number of proteins containing polyP, PPG or other stalling motifs is higher in eukaryotes and this could explain the fact that eIF5A is essential while bacterial EF-P is not.
Many questions about EF-P, still remain open, and to answer these questions further functional and structural studies are needed.

1.4 The archaeal translation factor aIF5A

1.4.1 Structural features

Archaea contains an eIF5A homolog protein and EF-P ortholog, termed archaeal initiation factor 5A (aIF5A), that shows strong structural and functional similarity to the eukaryotic and bacterial factors. This protein was isolated in 1992 from the thermophilic archaeabacterium *Sulfolobus acidocaldarius* as a small protein of 15 kDa containing a specific post-translational modification of a lysine residue, called hypusine, as described for eukaryotes (Bartig *et al,* 1992).
Concerning the structure, the crystal analysis of aIF5A from hyperthermophililc archaebacterium *Pyrococcus horikoshii* highlights the resemblance of the archaeal protein to the eukaryatic one.
Pho-aIF5A is predominantly composed of ß-strands and comprises two distinct domains (Fig.12A), an N-terminal domain which includes residues 1-69 and has a SH3-like barrel, consisting of a 3_{10}-helix (α1) and a six-ß-stranded (ß1→ ß6) anti-parallel ß-sheets. This

domain also includes a long hairpin loop called L1, that comprises residues 33-41, which hold a specific lysine (Lys37), that is supposed to be modified into hypusine, as it happens in eukaryotes. The C-terminal domain is composed of residues 72-138 and has an OB-fold (oligonucleotide binding fold) motif, comprising two short α-helices (α2- α3) and five stranded anti parallel ß-sheets (ß7→ ß11). These two domains are connected by a short linker of only two residues (70-71). Moreover, between the N- and C-terminal domain the basic residues create a positively charged surface, and also the L1 hairpin with hypusine modification is positive. On the contrary C-domain is mostly negatively charged and these difference in electrostatic potential of the protein led to the speculation that aIF5A can act as a bimodular protein capable of interacting with nucleic acid (especially RNA molecules) through his positive portion and conversely the C-terminal domain most negatively charged interact with proteins (Yao *et al*, 2003).

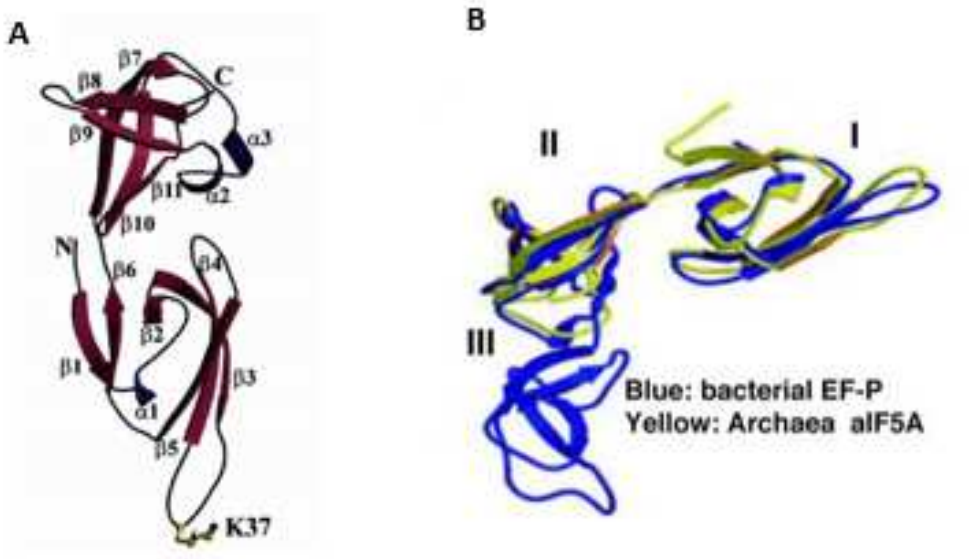


Fig. 12 (A) Representation of translation initiation factor 5A from P. horikoshii (Yao et al, 2003) (B) Comparison of crystal structure of aIF5A with EF-P (Park et al, 2010).

Comparing them to EF-P, eIF5A and aIF5A seem to share more similarity since both are lacking an additional third domain present in EF-P and have the same post translational modification (Figure 13A and B), though eukaryotic protein has an alpha-helix in C-terminal domain not present in aIF5A.

But if we consider the amino acids sequence between aIF5A and domain I and II of EF-P, identity between aIF5A/EF-P is greater than between eIF5A/EF-P (Aoki *et al*, 2008). The archaeal and bacterial structures are superimposable as shown in figure 12B, while eukaryal and bacterial proteins have similar overall shape, as eIF5A is a dimer in solution and resemble a L-shape protein exactly like EF-P (Figure 13 C and D). These structural similarities suggest that there are probably a structural and functional conservation of these translation factors throughout the evolution.

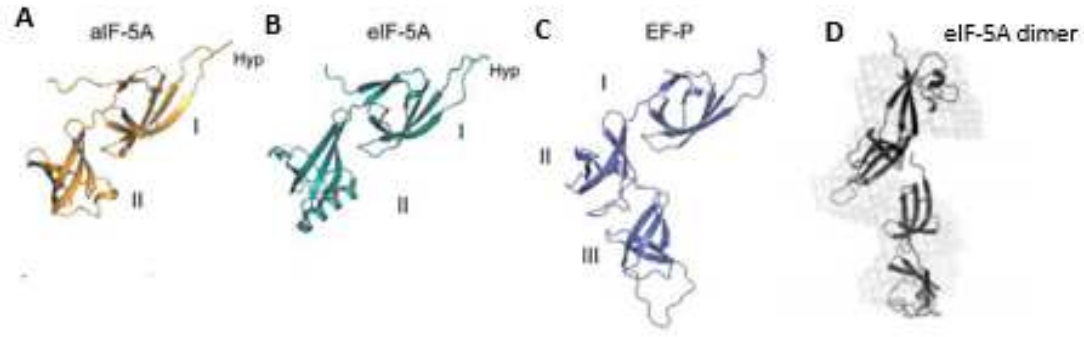


Fig.13 Structural insights into eIF5A, aIF5A and EF-P (Lassak et al, 2016) Comparison of crystal structure of (A) P. horikoshii aIF5A (PDB ID code 1IZ6), (B) human eIF5A (PDB ID code 3CPF) and (C) T. thernophilus EF-P (PDB ID code 1UEB). (D) Dimerization model of $eIF5A^{HYp}$ in solution from S. cerevisiae obtained by SAXS data (Dias *et al*, 2013).

1.4.2 Post-translational modification of aIF5A

Very little is known about the translation factor aIF5A in the archaeal kingdom. Like eIF5A, archaeal aIF5A is essential: the gene was found to be essential in *Haloferax volcanii* (Gäbel *et al,* 2013), in *Sulfolobus acidocaldarius* (La Teana and Albers, unpublished data) and more recently in *Sulfolobus islandicus* (Zhang *et al*, 2018).

Hypusine was identified since 1989 in some aerobic archaea such as *Sulfolobus acidocaldarius, Halobacterium cutirubrum* and *Thermoplasma acidophilum* (Schumann *et al*, 1990).

A year later a subsequent study revealed that the distribution of hypusine in the kingdom is heterogenous. Several species, strictly anaerobic thermophilic in the archaeal kingdom, contain deoxyhypusine instead of hypusine, the intermediate of eIF5A in eukaryotic process, such as *Thermoproteus tenax, Pyrodietium oecultum* and *Desulfurolobus ambivalens* (Bartig *et al*, 1990). Interesting, Thermoproteales *sp*. own also hypusine, suggesting that deoxyhypusine could be an intermediate of hypusine modification as in eukaryotes, but probably through an alternative hydroxylation pathway, since these anaerobic organisms are unable to perform an oxygenation reaction.

Although the mechanism of hypusine modification is widely studied in eukaryotes, the hypusination pathway and the role of aIF5A in Archaea are poorly understood. aIF5A is present in all archaeal genomes sequenced so far and also the DHS homologs are present in all sequenced archaeal genomes but no DOHH homologs has been identified in any archaeal genomes or proteomes to date (Park *et al*, 2006), raising a question about how hypusine residue is synthesized in certain archaeal family like Sulfolobaceae. Archaeal DHS is found to be essential as the eukaryal counterpart, because growth inhibition by N^1-Guanyl-1,7-diaminoheptane (GC7), an efficient DHS-inhibitor in eukaryotes, was demostrated in four archaeal species (*S. acidocaldarius, S. solfataricus, Halobacterium haloboum and Haloferax mediterranei*) (Jansson *et al*, 2000). These findings suggest that the first step of modification in Archaea is conserved and deoxyhypusine could be synthesized by similar mechanism as described in eukaryotes. GC7 competes with spermidine for the binding to DHS in eukaryotes; this compound associates with the enzyme at the same site of spermidine binding (Jakus *et al*, 1993), thus preventing the transfer of N-butilic-group of spermidine to lysine. The authors suggested that the presence of unmodified aIF5A, unable to synthetize a subset of proteins involved in cell cycle progression, lead to arrest of growth of these organisms at the end of the D period (G2).

A recent analysis of euryarchaeota *H. volcanii* aIF5A reveals that the protein is deoxyhypusinated and provides important information about the mechanism of this type of modification (Prunetti *et al*, 2016). First of all, a mass spectrometry analysis of *Hfx*-aIF5A proved for the first time that is modified with deoxyhypusine at the conserved lysine 36 (in 99% cases) and no hypusine is detected. Secondly, the inefficiency of GC7 on the *H. volcanii* growth and the absence of intracellular spermidine in this organism, suggest a different pathway and lead researchers to hypotesize a model in with agmatine, an essential polyamine

mainly present in halophiles, could be the substrate for the deoxyhypusine modification, differing from the canonical eukaryotic pathway.

Moreover, they characterized for the first time the archaeal DHS enzyme as a tetramer complex of 148 kDa, finding in agreement with the structural information of eukaryotic DHS.

However, this alternative modification reaction could not be a common feature of all Archaea or all narrow group to which *Haloferax* belongs, as DHS activity of another euryarchaea was explored as a control. In fact, Prunetti *et al.* showed for the first time an *in vitro* activity of *T. kodakarensis* DHS similar to eukaryotic enzyme: the transfer of the 4-aminobutyl group from spermidine to *Tko*-aIF5A. The modified aIF5A was analyzed by mass spectrometry demostrating the presence of deoxyhypusine residue on the lysine 42.

These findings pave the way for futher studies, necessary to elucidate hypusination pathway of translational factor aIF5A in Archaea, that could change from organism to organism.

As mentioned at the beginning, the function of aIF5A in Archaea is still poorly characterized and despite this protein seems to be similar to the eukaryal one, an involvment in protein synthesis has never been demostrated yet.

Wagner and coworkers showed for the first time an involvment of aIF5A from *Halobacterium sp.* in RNA-mediated processes, showing an RNA cleavage activity, which has not been reported for eukaryal eIF5A, and an RNA binding capacity. This latter activity required the hypusinated form of the protein, as hypusine residue is necessary to stabilize the RNA-aIF5A complex. Conversely, RNA degradation mediated by aIF5A does not require the hypusination of the protein, and the author indicated that this cleavage preferentially occurrs between adenine and cytosine nucleobases, within single strand regions (Wagner *et al*, 2007). Anyway, how these two activities are regulated is still obscure and further work is necessary to clarify these functions.

2. Materials and Methods

2.1 Purification of recombinant N-His-aIF5A from *E. coli*

The recombinant protein for *Sulfolobus solfataricus* aIF5A (ORF 0970) was expressed in *E. coli* and purified with a protocol described in Bassani, PhD thesis (2017).

Briefly, aIF5A gene was inserted into the plasmid pETM11 (kindly provided by Dr. Roberto Spurio, University of Camerino), to obtain a construct encoding a protein with six histidine residues at its N-terminal end. The recombinant plasmid was used to transform *E. coli* Rosetta (DE3) /pLysS cells. The resulting cells were grown at 37°C until they reached an optical density (OD) of 0.7 at 600 nm, then the production of the recombinant fusion protein was induced by addition of 0.5 mM IPTG (Isopropil-β-D-1-tiogalattopiranoside). After 3 hours cells were collected, resuspended in 20 ml of cold lysis buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 15 mM Imidazole, 1 mM PMSF (Phenylmethanesulfonyl fluoride), 10 mM ß-mercaptoethanol, 0.1% Triton X-100, 25 µg/ml lysozyme) and incubated for 30 min on ice. After lysis by sonication, the cell lysate was centrifuged at 25000 g for 30 min at 4°C and the fusion protein purified by affinity chromatography on Ni-NTA Agarose resin (Quiagen). Binding of the N-His-aIF5A to the resin was carried out overnight on a rotating wheel at 4°C. The beads were washed with 30 ml of Wash Buffer (50 mM Tris HCl pH 7.4, 500 mM NaCl, 40 mM Imidazole) and eluted in 2 ml Elution Buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 250 mM Imidazole). The eluate was dialyzed overnight at 4°C against Dialysis Buffer (50 mM Tris HCl pH 7.4, 150 mM KCl, 5% glycerol).

The concentration of aIF5A was determined by Bradford assay and the purity was assessed by SDS-PAGE (Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis), followed by Coomassie-blu staining. Aliquots of the protein were stored at -80°C.

2.2 Purification of recombinant protein aIF5A-C-His from *S. solfataricus*

Construction of plasmids for transformation of *S. solfataricus* PH1-16 cells is described in Bassani, PhD thesis (2017).

PH1-16 (pMJ05(ptf55)-aIF5A-C-His) strain was grown at 75°C in Brock's medium (composed of Brock's salts and supplemented with 0.2% NZamine, 0.2% sucrose, pH 3.0). At an $A_{600\ nm}$ of 0.8, the cells were harvested and lysed by sonication. Fusion proteins were

purified by affinity chromatography on Ni-NTA resin, as described for N-His-aIF5A from *E. coli.*

2.3 Purification of native aIF5A from *S. solfataricus*

Purification of native aIF5A was carried out essentially as described in Bartig *et al* (1992), with some modifications. Affinity chromatography was performed at 4°C, all other purification steps were done at room temperature. The post-ribosomal supernatant (S100) obtained form 20 g of cell paste was subjected to affinity chromatography on Cibacron blue Sepharose 6-fast flow (GE Healthcare, 50 ml bed volume, 25 x 2.5cm, 20 ml/h), equilibrated in Buffer A (20 mM Tris HCl pH 7.5, 0.1 mM EDTA, 150 mM KCl). The column was developed stepwise with Buffer A, Buffer A containing 6 mM spermine, again Buffer A, and finally Buffer A containing 1.5 M NaCl. 10 ml fractions from each step were collected and 10 µl from each fraction were tested by Dot-Blot (Biorad) using polyclonal anti-aIF5A antibody (Eurogentec), available in our lab. aIF5A positive fractions were pooled and concentrated using a Centricon 3KDa Amicon® Ultra centrifugal filter device. The following two chromatographic steps were performed using a FPLC Akta Purifier apparatus (GE Pharmacia) at room temperature. The concentrated protein pool was subjected to gel filtration chromatography on a Superose 12 column (GE Healthcare, 24 ml bed volume, 30 x 1.0 cm, 0.5 ml/min), equilibrated in Buffer A. Fractions of 0.5 ml were eluted with the same buffer and 2 µl from each fraction were tested by Dot-blot. aIF5A positive fractions were concentrated using Centricon 3 KDa (Amicon) and the buffer was changed to Buffer B (20 mM Tris HCl pH 9, 0.1 mM EDTA, 150 mM KCl). The sample was loaded on a Mono Q HR 5/5 column Pharmacia (1ml bed volume, 0.5 ml/min), equilibrated with Buffer B, and eluted employing a linear gradient of 0.05-1 M KCl. 0.5ml fractions were collected and 2 µl from each fraction were analyzed by Dot-blot. aIF5A positive fractions were concentrated and dialysed against aIF5A buffer (50 mM Tris HCl pH 7.5, 150 mM KCl).

Bradford analysis and SDS-PAGE were performed to determine protein concentration while Western Blot with anti-hypusine (Merk Millipore) antibody and LC-MSMS analysis were performed to identify the nature of protein modification.

2.4 Liquid chromatography – mass spectrometry (LC-MSMS analysis)

The purified native aIF5A was loaded on a Mini-PROTEAN TGX Stain-Free Precast Gel (BioRad), and the protein band was excised and analyzed by LC–MS/MS analysis. The gel band was digested with trypsin as described (Mair *et al*, 2015). After digestion the peptide solution was desalted on a custom-made C18 stagetip (Rappsilber *et al*, 2007). Tryptic digests were separated on an Ultimate 3000 RSLC nano-flow chromatography system (Thermo Fisher Scientific), using a pre-column for sample loading (PepMapAcclaim C18, 2 cm × 0.1 mm, 5 μm) and a C18 analytical column (PepMapAcclaim C18, 50 cm × 0.75 mm, 2 μm, Dionex-Thermo-Fisher Scientific), applying a linear gradient from 2 to 35% solvent B (80% acetonitrile, 0.1% formic acid; solvent A 0.1% formic acid) at a flow rate of 230 nl min−1 for 60 min. The eluting peptides were analyzed on a Q Exactive HFX Orbitrap mass spectrometer, equipped with a Proxeon nanospray source (Thermo Fisher Scientific). The data-dependent mode survey scans were obtained in a mass range of 375–1500 m/z with lock mass on, at a resolution of 60.000 at 200 m/z and an AGC target value of 3E6. The 10 most intense ions were selected with an isolation width of 1.6 Da, fragmented in the HCD cell at 27% collision energy, and the spectra were recorded at a target value of 1E5 and a resolution of 30000. Peptides with a charge of +1 were excluded from fragmentation, the peptide match and exclude isotope features were enabled and selected precursors were dynamically excluded from repeated sampling for 15 s. The raw data were processed with MaxQuant software package (version 1.6.0.16, http://www.maxqu ant.org) (Cox and Mann 2008) by searching against the sequence of aIF5A in the background of the *Sulfolobus solfataricus* Uniprot (http://www.unipr ot.org) and sequences of common contaminants, with tryptic specificity allowing 2 missed cleavages. Carbamidomethylation was set as fixed modification, oxidation of methionine, N-terminal protein acetylation, hypusine and deoxyhypusine on lysines as variable modifications, all other parameters were set to default. Results were filtered at a protein and peptide false discovery rate of 1%. The protein list was further filtered for a minimum of 2 unique and razor peptides. Peptide hits returned by MaxQuant were manually validated.

2.5 Purification of N-His-DHS and DHS C-His from *E. coli*

The gene for *S. solfataricus* Deoxyhypusine synthase (DHS) (ORF Sso0967) was amplified by PCR using 100 ng genomic DNA of *S. solfataricus* P2, with two pairs of oligonucleotides for cloning the gene in two different plasmids:

1. forward 5'-GCGGCCATGGTAAATAGAGAGGAC-3' (NcoI restriction site)
2. reverse 5'-CCGGGATCCTTAGCTTAATAAAGACG-3' (BamHI restriction site)
3. forvard 5'-AAAAGCATGCGCATAAATAGAGAGGACTTGTTAAAAAACCC-3' (SphI restriction site)
4. reverse 5'-AAAAGGATCCGCTTAATAAAGACGCGGCCAAAATAGG-3' (BamHI restriction site)

ORF Sso0967 amplified with primers 1 and 2 was cloned in pETM11 for expression of N-terminal His-tagged DHS, while primers 3 and 4 were used for cloning of aDHS in the pQE-70 plasmid (Quiagen) which adds a C-terminal His-tag to the recombinant protein.
N-His-aDHS and aDHS-C-His were expressed in *E. coli* BL21(DE3) and purified with the same protocol described for aIF5A.

2.6 Small-angle X-ray scattering (SAXS) experiments

The recombinant Sso N-His-aIF5A and DHS-C-His proteins were produced from *E. coli* and purified by affinity chromatography as described above. The two proteins were, subsequently, subjected to SAXS experiments in order to obtain structural details about the overhall shape of the proteins and their conformation on solution at high temperature, and also to verify a formation of a stable complex between aIF5A and DHS.
The measurements were carried out at the Diamond Light Source (Harwell Science & Innovation Campus, Didcot, UK) synchrotron SAXS beamline BL21 with monochromatic beam. We have measured SAXS curves of the aIF5A protein, in 20 mM Tris HCl pH 7.7 and 2 % glycerol, at five different protein concentration (0.5, 1, 2, 5 and 10 mg/ml) with 2 different KCl amounts (60 and 1000 mM). Data have been recorded in the capillary heater at step of 5°C from 20°C to 63°C, the maximum temperature achievable by the apparatus. The range of the scattering vector moduls q ranged from 0.01 to 0.37 Å-1. Sso DHS protein in 20 mM Tris HCl pH 6.8, 120 mM KCl, 2% glycerol, were measured at 63°C at a concentration of 4 mg/ml, alone and supplemented with 2 mM NAD^+ and 2 mM spemidine.

To assessed formation of a complex between aIF5A and DHS, scattering was recorded with the two proteins mixed in a molar ratio of 4:1, in the absence and in the presence of all the other components of the modification reaction, 2 mM NAD^+ and 2 mM spermidine. Buffer scattering was recorded before each sample scattering.

All protein preparations were centrifuged for 10 min at 14000 g before the measurements, to remove aggregates or particles, and stored at 4°C. During the analysis the samples were kept at room temperature.

2.7 SAXS data analysis

All analysis of SAXS data were performed in collaboration with the biophysics group of Polytechnic University of Marche (Ancona), coordinated by Prof. Francesco Spinozzi.

The radius of gyration (Rg) and molecular weight of the aIF5A protein was determined by the Guinier equation and with a model that combines Guinier and Porod analysis (fit Guinier-Porod). For curves construction the *Gnuplot* program was used, editing the *scripts* with *Emacs* program.

A Kratky plot is obtained by plotting scattering intensity as $I(q)\times q^2$ vs. This representation divides out the decay of the scattering, making certain other features more evident.

The QUAFIT software was used to determine the tertiary structure of the protein (Spinozzi and Beltramini, 2012), from the analysis of SAXS curve of aIF5A 10 mg/ml, 60 mM KCl, 63°C.

2.8 Size-exclusion molecular chromatography

2.8.1 Size-exclusion molecular chromatography of aIF5A

The recombinat Sso proteins N-His-aIF5A from *E. coli* and aIF5A-C-His from *S. solfataricus,* were loaded on a 24 ml Superdex 75 10/300 GL Tricorn column (GE Life Science). The molecular mass standard proteins (Sigma-Aldrich) BSA (Bovine serum albumin) (67KDa) and myoglobin (17KDa), each at a concentration of 1 mg/ml, were used to calibrate the column. The void volume of column was calculated using Blue Dextran 2000. The chromatographic buffer was subjected to filtration and contained 50 mM Tris HCl pH 7.4, 150 mM KCl. The column was equilibrated with 3 volumes of buffer, than 0.5 mg/ml

of different aIF5A preparations, pre-heated at 65°C for 10 min and centrifuged at 10000g for 10 min, were separated on a gel-filtration column using 1.5 volumes of buffer (35 ml), with 0.4 ml/min of flow rate. Elution fractions (500 µl) were collected and those containing protein selected according to their absorbance at 280 or 254 nm after which the proteins were analyzed by SDS-PAGE and visualized with Western Blot, using anti-aIF5A antibodies.

2.8.2 Size-exclusion molecular chromatography of DHS-aIF5A complex

A 3 ml Superdex 200 increase 5/150 GL Tricorn column (GE Life Science) was used to identify complex formation between aIF5A and DHS.

N-His-aIF5A and DHS-C-His were purified from *E. coli* as described above. Both recombinant proteins were used in 50 µl enzymatic assays containing a reaction buffer (20 mM Tris HCl pH 7.4, 150 mM KCl), 50 µg of DHS-C-His, 25 µg of N-His-aIF5A, 2 mM NAD^+, 2 mM spermidine. Reactions were also performed with aIF5A alone (20 µg), DHS alone (20 µg), DHS (25 µg) with 2 mM NAD^+ and 2 mM spermidine in the same buffer. All reactions were incubated at 65°C for 30 min, and then loaded onto a gel filtration column, set at 65°C and pre-equilibrated with 9 ml of reaction buffer.

BSA (67 kDa), lactate-dehydrogenase (140 kDa) and myoglobin (17 kDa) were used as mass standard proteins, each at a concentration of 1 mg/ml. The void volume of the column was calculated using Blue Dextran 2000.

Elutions were carried out with 8ml of reaction buffer at 0.4 ml/min flow rate.

Elution fractions (400 µl) were collected and analyzed by SDS-PAGE followed by Western Blot, using anti-histidine antibodies, to appreciate complex formation between aIF5A and DHS.

2.9 Non-denaturing gel elecrophoresis

Mixtures of purified recombinant *Sso* proteins N-His-aIF5A (5 µg, 300 pmol) and deoxyhypusine synthase DHS-C-His (10 µg, 80 pmol) from *E. coli* in a molar ratio of 1:4 (enzyme/substrate), in the presence or absence of 1 mM NAD^+ and 1 mM spermidine, were incubated in 200 mM glycine/ NaOH pH 8.2 at 65°C for 30 min. After addition of native gel sample buffer 6x (187.5 mM Tris HCl pH 8.8, 75% glycerol, 3% bromophenolblue), the

proteins were separated under non-denaturing conditions on 12% polyacrilamide gel using following receipt:

- for a 10 ml native PAGE separating gel

Acylamide percentage	6%	8%	10%	12%	15%
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	2ml	2.6ml	3.4ml	4ml	5ml
0.375M Tris-HCl(pH=8.8)	7.89ml	7.29ml	6.49ml	5.89ml	4.89ml
*10% (w/v) ammonium persulfate (AP)	100µl	100µl	100µl	100µl	100µl
*TEMED	10µl	10µl	10µl	10µl	10µl

- for a 5 ml native PAGE stacking gel

0.375 M Tris-HCl pH=8.8	4.275ml
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	0.67ml
*10% (w/v) ammonium persulfate (AP)	0.05ml
*TEMED	5µl

Gel was run in 25 mM Tris base, 192 mM glycine running buffer at 120 V for 2 hours, at 4°C.

Proteins separation was determined through Comassie Brilliant Blue staining. To actually prove complex formation between aIF5A and DHS, the corresponding band of the complex, detected by staining with Comassie, was excised from the native gel, cut into fine pieces, incubated with an equal volume of Sample Buffer 1X (62.5 mM Tris HCl pH 6.8, 2% SDS, 10% Glycerol, 5% ß-mercaptoethanol, 0.002% bromophenolblue) and boiled for 10 min at 98°C. Sample was loaded on a 15% SDS-PAGE and the two band of 5A and DHS were visualized by Comassie Blue staining.

2.10 *In vitro* hypusination assay

The Deoxyhypusine Synthase activity was tested *in vitro* using different amount of the recombinant protein, purified from *E. coli*, after pre-incubation for 10 min at 65°C. 200 and

1200 pmol of DHS-C-His were incubated in presence of 400 pmol of recombinant N-His-aIF5A purified from *E. coli*, 2 mM spermidine, 2 mM NAD^+, 2 mM $MgCl_2$, 50 mM glycine/NaOH pH 9.4, 150 mM KCl and 1 mM DTT in a final volume of 30 µl. The reaction mixture was incubated for 2 h at 65°C. An eventual increase in mass of N-His-aIF5A after the reaction was monitored by LC-MS. The recombinat protein N-His-aIF5A produced in *E. coli* and, therefore, unmodified, was analyzed as a control.

High performance liquid chromatography was performed on a Dionex Ultimate 3000 HPLC (Thermo Fisher Scientific) system configured with the Chromeleon 6.0 software (Thermo Fisher Scientific).

Proteins were reduced in 100 mM DTT for 30 minutes at room temperature and then separated on an Aeris Widepore C4 column (3.6 µm particle size, dimensions 2.1 x 150 mm, Phenomenex) running a six minutes step gradient from 10% up to 70% acetonitrile in 0.1% formic acid.

The working temperature was set at 50°C and the flow rate at 300 µl/min. The LC-system was coupled online to the quadrupole-time of flight-mass spectrometer Synapt G2-Si (Waters), operated via the MassLynx V 4.1 software package, using a Z Spray ESI source (Waters). Mass spectra were acquired in the m/z range from 500-2000, at a scan rate of 1 sec and the mass spectrometer was calibrated with a MS spectrum of [Glu1] -Fibrinopeptide B human (Glu-Fib) solution.

Acquired data were analyzed with the MaxEnt algorithm to reconstruct the uncharged average protein mass.

2.11 Preparation of *S. solfataricus* cell lysate

Sulfolobus solfataricus P2 culture was aerobically grown in Brock's medium (Brock *et al,* 1972), at 75°C and pH 3, with supplement of 0.2% NZ-amine and 0.2% sucrose.

Cells at exponential phase of growth (0.8 OD_{600}) were harvested and pelleted. *S. solfataricus* cells extract was prepared according to Benelli and Londei (2007) starting from 20 g (wet weight from 20 L of culture) of frozen cells. These were distrupted by grinding, with a sterile, cold mortar on ice, adding 40 g of alumina (2 g alumina each g of cells). Cells were resuspended in 20 ml of Extraction Buffer (Tris-HCl pH 7.4 20 mM, MgAc 10 mM, NH_4Cl 40 mM, ß-mercaptoethanol 6 mM), and centrifuged at 15300 x g for 30 min at 4°C.

The supernatant was centrifuged at 15300 x g for 45 min at 4°C, and crude cell extract S30 was collected. S30 was than ultracentrifuged with Beckman Type 55.2 Ti rotor at 30000 rpm (100000 g) for 16 h at 4°C. Resulting pellet was ribosomal fraction and the supernatant was cytoplamic fraction, called S100, rich in proteins and cytoplasmatic factor.

S30 and S100 concentration was determined by Bradford.

2.12 Analysis of aIF5A levels under different growth conditions

The *S. solfataricus* cells were aerobically cultivated at 75°C pH 3, as previously described. Cell growth was monitored measuring A_{600nm} until late exponential growth phase (34 h of growth). When culture reached different densities, 15 ml of culture at 0.2 OD (8h), 0.4 OD (18h), 0.8 OD (24h), 1.3 OD (26h), 1.8 OD (30h) and 1.82 (34h), were harvested by centrifugation at 4000 rpm for 20 min at 4°C. Pellets were resuspended with Extraction

Buffer (20 mM Tris HCl pH 7.4, 20 mM MgAc, 40 mM NH_4Cl, 1 mM DTT) and lysed by four cycles of incubation at 37°C/-80°C for 5 min.

Lysates were clarified through centrifugation at 15000 rpm for 30 min at 4°C.

The total protein concentration in the supernatants was measured by the Bradford assay.

Cold and heat-shock treatments were performed as follows. Cells grown at 75°C to the exponential phase were transferred to either 60°C or 90°C for 30 min. Cells were than harvested and the lysates prepared as described above.

In all cases, 20 µg of total extract proteins was fractionated by SDS-PAGE and the presence of aIF5A was revealed by western blot, probed with anti-aIF5A.

2.13 Preparation of salt-washed ribosomes

For detailed determination of the intracellular localization of aIF5A, the ribosome pellet (termed crude ribosomes) was resuspended in a high-salt buffer named Suspension Buffer (20 mM Tris HCl pH 7.4, 10 mM MgAc 500 mM NH_4Cl, ß-mercaptoethanol 6 mM). The ribosome suspension was clarified through a centrifugation at 15000 g for 20 min at 4°C. The supernatant was then layered onto a 20 ml cushion of 18% sucrose in Suspension Buffer and centrifuged in Beckman Type 55.2 Ti rotor at 30000 rpm (100000 g) for 16 h at 4°C. The ribosome pellets (termed "saltwashed ribosomes") were lastly resuspended with Extraction Buffer and together with the supernatant termed "high salt wash" HSW, were stored at -80°C.

Proteins concentration of crude ribosome, saltwashed ribosome and HSW was determined by Bradford assay. These fractions were analyzed by Western Blot, with anti-aIF5A antibody to determine the sub-cellular localization of the protein.

2.14 Isolation of ribosomal subunits

Ribosomes were dissociated and 50S and 30S subunits separated by centrifugation of the salt-washed particles on cold 15 ml linear 10-30% (w/v) sucrose density gradients in sterile Extraction Buffer. 40 OD_{260} units of salt-washed ribosomes were stratified on each gradient and were ultracentrifuged in Beckman SW28 rotor operated at 23000 rpm for 16 h at 4°C. Fractions corresponding to the 50S and 30S peaks of A_{260} were separately pooled and centrifuged with Beckman Type 55.2 Ti rotor at 35000 rpm (100000 g) for 18 h at 4°C.

Finally, pellets of 30S and 50S were resuspended with Extraction Buffer and concentration was determined from absorbance at 260 nm, knowing that 1 OD of 50S correspond to 60 pmol, and 1 OD of 30S is 70 pmol of ribosome subunits.
These ribosomal subunits were used for specific aIF5A binding experiments with 30S or 50S of *S. solfataricus*.

2.15 Interaction of aIF5A with ribosomes

The interaction of N-His-aIF5A (*E. coli*) and aIF5A-C-His (Sso) proteins to ribosomal subunits was investigated by mixing increasing concentration of purified *S. solfataricus* ribosomes (20, 40, 80 pmol of 30S and 50S) with 30 pmol of recombinant His-tagged aIF5A in 60 µl (final volume) of binding buffer (10 mM KCl, 20 mM $MgCl_2$, 20 mM Tris HCl pH 7.4, 6 mM ß-mercaptoethanol). Reactions were incubated at 65°C for 10 min and were stratified on 30 µl of 10% sucrose cushion, followed by a centrifugation at 100000 rpm 4°C for 1 h (S100AT3 rotor). 30 µl of supernatant were withdrawn and separated on 15% SDS-polyacylamide gel. The presence of aIF5A was probed with anti-aIF5A antibodies.

2.16 Western Blot analysis

Standard protocol (Laemmli,1970) was used for SDS-PAGE, Sample Buffer 2x (0.125 M Tris HCl pH 6.8, 4% SDS, 20% Glycerol, 10% ß-mercaptoethanol, 0.004% bromophenolblue) was added to the samples, boiled for 10 min at 98°C, and loaded onto 15% polyacrylamide gels. After electrophoresis samples were transferred onto a 0.2 µm nitrocellulose membrane (GE Healthcare) using a Semi-dry blotting apparatus (Trans-blot SD Semi-Dry Transfer Cell, BioRad). Protein transfer was performed at 15 V for 20 min in Transfer Buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol). After blocking non-specific binding with 5% nonfat milk in TBST Buffer (20 mM Tris Base, 150 mM NaCl, 0,1% Tween 20) for at least 1 h at 37°C, the blots were probed with anti-aIF5A (Eurogentec), in a dilution 1:10000 in TBST-Buffer added to 5% nonfat dried milk powder (Euroclone), with anti-hypusine antibody (Millipore) 1:2000 in TBST-BSA 3% or with anti-histidine (Thermo Scientific) 1:6000 in TBST-BSA 1% overnight at 4°C. Membranes were washed three times with TBST. A secondary goat anti-rabbit IgG-HRP antibody (Santa Cruz Biotechnology) was used for anti-aIF5A and anti-hypusine at a 1:10000 dilution in TBST-

BSA 3%, while goat anti-mouse IgG-HRP (Santa Cruz Biotechnology) was used for anti-histidine, for 1h at 37°C. The membranes were washed with TBST for four times and subsequently were developed with the enhanced chemiluminescence substrate for Western Blotting LiteAblot PLUS (Euroclone), according to the manufacturer's instructions. The images were visualized with BioRad ChemiDoc Imaging system.

2.17 Sucrose gradients

One gram (wet weight) of wild-type cells was lysed in buffer containing 20 mM Triethanolamine (TEA)-HCl pH 7.5, 10mM MgAc, 40 mM NH_4Cl, 6 mM ß-mercaptoethanol. The cell lysate was centrifuged as described above and 500 µg of total protein were pre-incubated at 70°C for 10 min and then were "programmed" for translation as described in Condo *et al* (1999). The reaction contained 10 mM KCl, 20 mM TEA, 20 mM MgCl2, 2 mM ATP, 1 mM GTP, 1.5 µg of *S. solfataricus* tRNA, 0.1 mM amino acids mixture, 4 µg of an *in vitro* transcribed mRNA (Sso5866 mRNA), in a final volume of 100 µl. The programmated lysate was incubated at 70°C for 15 min to activate translation, and then fixed with 2% formaldehyde to stabilize 70S ribosomes which are easily dissociated in *S. solfataricus* (Londei *et al*, 1986). After incubation on ice for 1 h, the lysate was loaded on a 10-30 % linear sucrose density gradient in 20 mM Triethanolamine (TEA)-HCl pH 7.5, 10 mM MgAc, 40 mM NH_4Cl, 6 mM ß-mercaptoethanol. As control 500 µg of S30 lysate non programmed for protein synthesis was loaded on the same sucrose gradient, using 14 ml centrifuge tubes. After centrifugation at 23000 rpm (Beckman Optima L-90K Ultracentrifuge SW28 rotor) for 17 h at 4°C, 500 µl fractions were collected by continuously monitoring absorbance at 260 nm. Samples were precipitated with 10% Trichloroacetic acid (TCA), loaded on 15% SDS-polyacylamide gels and analyzed by Western blot analysis with anti-aIF5A antibodies and anti-aIF6 as a control (Benelli *et at*, 2009b).

2.18 Immunoprecipitation (IP) of native aIF5A from *Sulfolobus solfataricus* lysate and RNA extraction

The citoplasmatic fraction of *S. solfataricus* lysate (S100), prepared as described above, was used for immunoprecipitation of aIF5A.

80 mg of S100 total protein concentration and 80 µl of anti-aIF5A antibody were incubated together on a rotating wheel overnight at 4°C to allow specific binding of native aIF5A to his antibody. Pre-immune serum (80 µl) was used as a control for anti-aIF5A antibody.

Dynabeads Protein G beads (Thermo Fisher Scientific) were equilibrate with 3 x 1 ml of Equilibration/ Wash Buffer (50 mM Tris HCl pH 7.4, 100 mM NaCl, 10 mM MgAc, 0.1% Triton X-100), 50 µl were then added to each reaction and samples were incubated for 2 h at 4°C. Immunocomplexes were captured by magnetic device and washed with 3 x 1 ml of Wash Buffer. Unbound fractions and the last washes were collected and analyzed by Western Blot.

Proteins were eluted with 100 µl of SDS-sample buffer 1x, boiling beads at 95°C for 10 min on vigorous shaking and 20 µl of IP samples were analyzed by western blotting using anti-aIF5A antibody.

In order to isolate RNA molecules directly bound to native aIF5A protein, the other 80 µl of IP samples were subjected to RNA extraction with Phenol/Chloroform treatment to denaturate proteins, RNA precipitation by Ethanol and resuspension in 10 µl of DEPC water. Samples were treated with TURBO™ DNase (RNase-free, Thermo Fisher Scientific) and extraction with Phenol/Chloroform. Concentration of RNA was determined with Nanodrop 2000c spectrophotometer device.

To confirm complete degradation of chromosomal DNA, 16S PCR was performed with Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific) and then loaded into a 6% native polyacrilamide gel.

The primers used forward 5'-TTGGGATCGAGGGCTGAAAC-3' and reverse 5'-CTCACCCCTCTCCTACTCGG-3', amplified a short part of 16S rRNA gene (133 nt at the 3'-end).

	TEMPERATURE	TIME	CYCLES
Initial denaturation	98°C	5’	1x
Denaturation	98°C	30’’	30x
Annealing	55°C	30’’	
Elongation	72°C	30’’	
Final extension	72°C	10’	1x

Table 2. PCR program for the amplification of 16S rRNA

2.19 Reverse transcriptase-PCR assay

In a previous deep-sequencing analysis of RNAs co-purified with aIF5A, performed by Dr. Flavia Bassani and described in her PhD thesis (2017), a list of RNAs associated with recombinant protein obtained after affinity purification of lysate from Sso PH1-16 [pMJ05 (ptf55) -aIF5A-C-His], was identified. In order to validate this data an *in vivo* association of some of this RNA with the native protein was tested with RT-PCR (Reverse Transcriptase-Polymerase Chain Reaction).

Three of this RNAs were selected for high number of reads compared to the control (RNAs co-eluted in a mock purification, Sso PH1-16 [pMJ05 (ptf55-empty plasmid)]):

- Sso 2184 mRNA
- Sso 0910 mRNA
- Sso ncRNA 98

and one Sso mRNA not present in a list of co-purified mRNA was chosed as a control (Sso 2508sh mRNA) (Hasenöhrl *et al,* 2008).

For cDNA synthesis, 50 ng RNA template extracted after immunoprecipitation with anti-aIF5A antibodies and with pre-immune serum, was mixed with 100 ng of random primers (Promega), dNTP mix (10mM each). The mixture was then treated at 65°C for 5 min, followed by 1 min incubation on ice. The cDNA synthesis was performed with SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer’s instructions. The total mix in a final volume of 20 µl was incubated at 25°C for 5 min, followed by 50°C for 90 min. The reaction was then inactivated at 70°C for 15 min.

2 µl of cDNA derived from RNA co-immunoprecipitated with aIF5A and control (pre-immune serum), respectively, were used as a template for PCR amplification using pairs of primers that amplified an internal region of about 200 bp of cDNA sequences.

- Sso 2184 mRNA:
  Forward 5'-GAGCAATTATATATCTTGACGAGGTTGACACG-3'
  Reverse 5'-AGCATACTTCGAAAGAATAAATTTCAGTTGC-3'
- Sso 0910 mRNA:
  Forward 5'- CCGTTGCTAAAATATTAGGCGATTTGAAG -3'
  Reverse 5'-CTGGTAATAATTCCCTAACTTTAACTTCAGCC-3'
- Sso ncRNA 98:
  Forward 5'-TAATACGACTCACTATAGGGATCTTTGGTTTAGCATCTCTC-3'
  Reverse 5'-GTTAGAGAAAACGGAGAGG-3'
- Sso 2508sh mRNA:
  Forward 5′-AGATAATACGACTCACTATAGATGATTGTAGGATTTGCCGGAAAACT-3′
  Reverse 5′-GTTAGCCATTATCCCATCGACGTCAGCG-3′

The corresponding PCR products were 230 nt (mRNA 2184), 225 nt (mRNA 0910), 110 nt (ncRNA 98) and 253 nt (mRNA 2508sh) in lenght. To confirm total DNA depletion of RNA samples, the PCR reactions were performed with the same amount of RNA as used for RT-PCR reaction without reverse transcriptase.

	TEMPERATURE	TIME	CYCLES
Initial denaturation	98°C	10'	1x
Denaturation	98°C	30''	30x
Annealing	55°C	30''	
Elongation	72°C	45''	
Final extension	72°C	10'	1x

Table 3. PCR program for the amplification of Sso 2184, 0910, 2508sh mRNA and ncRNA 98

PCR products were loaded into 6% polyacrilammide native gel. The electrophoresis was carried out at 10 mA for 40 min, using Tris-Borate-EDTA (TBE) as running buffer.

2.20 *In vitro* transcription

PCR template for generation of Sso 2184 mRNA, 0910 mRNA and ncRNA 98 were prepared by using *S. solfataricus* genomic DNA as template and the following oligonucleotides (forward primers contain the T7 promoter sequence):

5'-
AGATAATACGACTCACTATAGGGTGCATGTGATAAGGGAAACATTGAAAGG-3' (Fwd 2184 mRNA),

5'-TAATTTTCTCCCACATCATCAAAATCACC-3' (Rv 2184 mRNA),

5'-
AGATAATACGACTCACTATAGGATGTTTGATAAGTTACCATTTATTTTTAACAACG-3' (Fwd 0910 mRNA),

5'- TTAACTCTCCACCGCTATTAGTCCCTTATTC -3' (Rv 0910 mRNA),

5'-TAATACGACTCACTATAGGGATCTTTGGTTTAGCATCTCTC-3' (Fwd nc98),

5'-GTTAGAGAAAACGGAGAGG-3' (Rv nc98).

Sso-2508sh mRNA was in vitro transcribed as described in Hasenöhrl *et al,* 2008.

The RNAs were synthesized in vitro using T7 RNA polymerase (Fermentas) for 3 h at 37°C, followed by TURBO™ DNase (RNase-free, Thermo Fisher Scientific) treatment for 30 min at 37°C. RNAs were then purified by Phenol/Chloroform extraction and resuspended in 50 µl of DEPC-water.

2.21 5'-end-radiolabelling of 2508sh mRNA and ncRNA 98

200 pmol of *in vitro* transcribed 2508sh mRNA and ncRNA98 were dephosphorylated with FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific) for 10 min at 37°C. 10 pmol of dephosphorylated RNAs were radiolabeled at the 5´-end with T4 Polynucleotide Kinase (Thermo Scientific) in the presence of 15 pmol of [α-32P] ATP (3000 Ci/mmol, Hartmann Analytic GmbH). Radiolabeled 2508sh RNA was purified through a ProbeQuant™ G-50 Micro Column (GE Healthcare) to remove unincorporated radiolabeled nucleotides and then phenol/chloroform extracted.

2.22 *In vitro* RNA degradation assay

RNAs *in vitro* transcribed as described above were first denatured for 5 min at 85°C and then incubated with Sso aIF5A proteins performing different reactions in a degradation buffer (10 mM HEPES pH 8, 40 mM KCl, 0.5% glycerol, 5 mM ß-mercaptoethanol, 5 mM $MgCl_2$ in DEPC-water). 2508sh, 2184 mRNA, 0910 mRNA and ncRNA 98 were the substrates for this cleavage assays performed by N-His-aIF5A produced in *E. coli* and aIF5A-C-His produced in Sso.

20 pmol of the different RNAs were incubated with 40 pmol of aIF5A (ratio RNA/5A 1:2), 80 pmol (ratio RNA/5A 1:4) and 160 pmol (ratio RNA/5A 1:8) at 65°C for 15 min. As a control of RNA stability 20 pmol of RNA without protein was also incubated at 65°C for 15 min. Afterwards, 160 pmol of RNA were incubated with 640 pmol of aIF5A (RNA/5A 1:4) in a degradation buffer for 30 min at 65°C in a total volume of 40 µl. Samples of 5 µl (20 pmol of RNA each) were withdrawn every 5 min for 30 min total. As a control 20 pmol of RNA without protein was also incubated at 65°C for 30 min.

An equal volume of RNA loading dye (Thermo Fisher Scientific) was added to every samples and, before loading on a denaturing (8 M urea) polyacrylamide 8% gel, were incubated at 85°C for 3 min. The electrophoresis was carried out at 20 mA for 30 min, using Tris-Borate-EDTA (TBE), and gels were staining for 10 min with 10 µg/ml of Ethidium Bromide.

The RNase activity of recombinant N-His-aIF5A produced in *E. coli* (6 µM) towards 2508sh RNA (10 nM 5´-end radiolabelled RNA and 1.5 µM cold RNA) was assayed between 0 and 24 min at 65 °C and every 2 min an aliquote was withdrawn. As a control 5´-end radiolabelled RNA was incubated for 24 min at 65 °C in the absence of recombinant N-His-aIF5A produced in *E. coli*. The samples were loaded on a 20% PAA-8M urea gel, which was analyzed with a Typhoon FLA 9500 PhosphorImager (GE Healthcare).

2.23 Electrophoretic mobility shift assay (EMSA)

5'-labelled-nc98 was incubated either with N-His-aIF5A (expressed in *E. coli*) or with hypusinated aIF5A-C-His (expressed in Sso). Increasing concentration of aIF5A proteins were incubated with 5'- labeled -nc98 in a molar ratio of 0:1, 10:1, 50:1, 100:1, 200:1, 400:1, 600:1, 800:1. The RNA and the corresponding proteins were incubated in a volume of 20 µl

containing 10 mM HEPES pH 8, 40 mM KCl, 0.5% glycerol, 5 mM ß-mercaptoethanol, 5 mM $MgCl_2$ in DEPC-water, at 65°C for 20 min. Then 2 µl 85% glycerol were added and the samples were loaded on a 8% polyacrylamide gel. The electrophoresis was carried out at 90 V at 4°C for 120 min, using cold TBE (Tris-Borate-EDTA) as running buffer. The gel was dried and the radioactive signals were detected with a Typhoon FLA 9500 PhosphorImager (GE Healthcare).

3. Results

The *S. solfataricus* gene predicted to encode aIF5A (ORF Sso0970) is most likely essential, but little is known about this gene/protein in this organism. To obtain a better understanding of aIF5A in Archaea and specifically in Sulfolobaceae we have performed different experiments in order to unveil first of all the post-translational modification of the protein, then try to characterize the structure and the function of the proteins involved in the hypusination pathway, the translation factor aIF5A and the first enzyme of modification DHS, both homologous between Archaea and Eukarya.

3.1 *Sulfolobus solfataricus* native aIF5A is hypusinated

Analyses based on amino acids composition data reported the presence of both hypusine and deoxyhypusine in Archaea. Hypusine was detected in several Crenarchaea like *Sulfolobus acidocaldarius, Pyrodictium occultum, Thermoproteus tenax* and *Acidianus ambivalens,* while deoxyhypusine was found in Euryarchaeota, as halobacteriales, methanogen, thermococcales and thermoplasmales (Bartig *et al*, 1992). The specific nature of the modification found in archaeal aIF5A was confirmed by mass-spectrometry only in *Haloferax volacanii*, where aIF5A is deoxyhypusinated at the conserved lysine 36 (Prunetti *et al*, 2016).

In order to determine whether the *S. solfataricus* aIF5A is modified by hypusine or deoxyhypusine, the native protein was isolated from Sso lysate. The native aIF5A protein was purified according to the protocol described for *S. acidocaldarius* (Bartig *et al*, 1992) with some modification, as described in Materials and Methods. Starting from 180 mg of total proteins in *S. solfataricus* P2 S100 lysate, 50 µg of native aIF5A were purified, in three different chromatographic steps. In the first step the lysate was loaded on the affinity chromatography column Cibacron blue Sepharose 6-fast flow, using polyamine spermine to induce elution of the protein. Fractions of 10 ml were collected (W= washes, E= elutions after the addition of spermine) and analyzed with Dot-Blot (BioRad) with anti-aIF5A antibody (Figure 14).

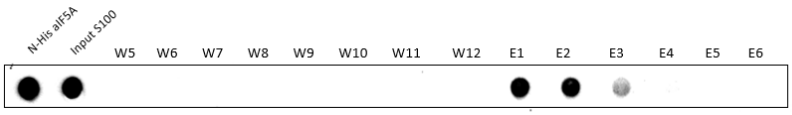


Fig.14 Dot-blot analysis of Cibacron blue Sepharose 6-fast flow fractions with anti-aIF5A

Fractions E1, E2, E3 were pooled, concentrated to 1.5 ml with Centricon 3kDa (Amicon) and loaded on a Superose 12 column, for the gel-filtration chromatographic step. Figure 15A shows the elution profile. 0.5 ml fractions were collected and those below the main A_{280} nm peak were immunodetected with anti-aIF5A using again Dot-Blot technique (Fig.15B). aIF5A-positive fractions (23,24,25) were pooled.

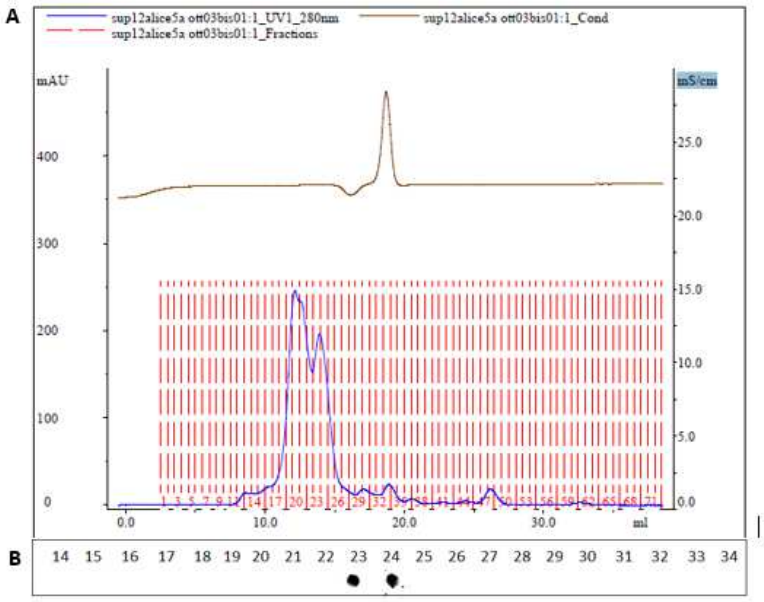


Fig.15 ***Gel filtration chromatography on Superose 12 of native aIF5A.*** (A) Elution profile of Cibacron positive-aIF5A pool fractions. (B) Analysis of Superose 12 fractions (14→34) by Dot-Blot with anti-aIF5A antibody.

The final purification step was performed on a Mono Q column, using a linear gradient of KCl concentration (50 mM → 1000 mM). 0.5 ml fractions were collected and analyzed by Dot-Blot. Figure 16 shows the elution profile (A) togheter with the Dot-blot result (B).

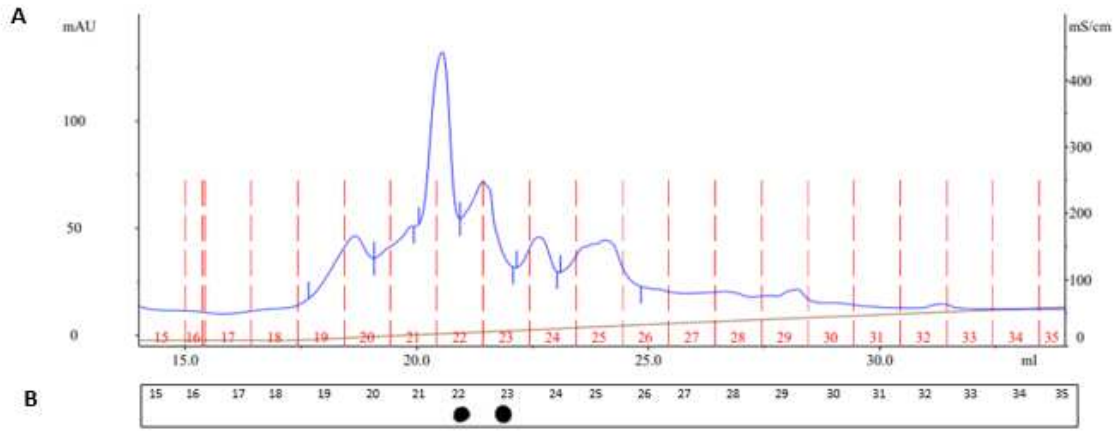


Fig.16 Anion-exchange chromatography of native aIF5A on Mono-Q. (A) Elution profile of MonoQ pool. Fractions of 0.5 ml were collected (B) Analysis of Mono Q fractions (14→35) by Dot-Blot with anti-aIF5A antibody.

Finally, the positive fractions (22, 23) were concentrated and protein concentration was determined by Bradford analysis. To determine the purity of aIF5A, an aliquote was loaded on SDS-PAGE gel and stained with Comassie Blue (Fig.17A). The major protein band is aIF5A having the expected molecular weight, however other proteins are also present in the sample, and to identify the native aIF5A protein, we performed a Western Blot analysis using a commercial anti-hypusine antibody. The antibody, which was raised against a synthetic peptide encompassing the hypusination site of human eIF5A, specifically recognized a single protein providing the first evidence that aIF5A from *S. solfataricus* is modified. It is reported that this antibody is able to recognize both deoxyhypusine and hypusine residue, therefore to unveil the nature of the modification, the corresponding band was excised from the gel, digested with trypsin, and analyzed by LC-MS/MS. Figure 17B shows the spectrum of the hypusinated peptide TGKHGSAKANVVAIGVFSGAK, revealing the hypusine modification at the conserved K36 residue of native aIF5A. The mass of the charged precursor is 2087.18 (1043.59*2), which corresponds exactly to the mass of the unmodified peptide (1999.30) and the hypusine amino acid (86.77). Moreover, it is important to mention that no deoxyhypusine residue was observed.

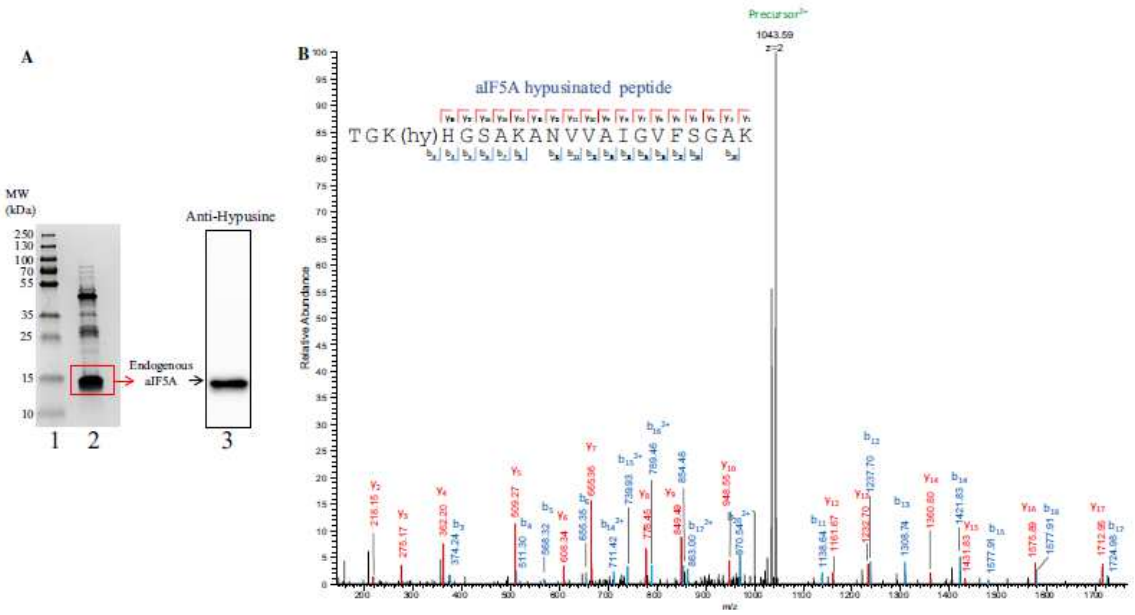


Fig.17 Native Sso-aIF5A is hypusinated. (A) Lane 1, molecular weight standards, lane 2, Comassie stained native aIF5A purified from *Sulfolobus solfataricus* P2. Lane 3, Western Blot analysis of native aIF5A probed with anti-hypusine antibodies. (B) MS spectrum of the charged precursor.

These results have given us important information: the native aIF5A in *Sulfolobus solfataricus* is hypusinated *in vivo*, as the eukaryotic eIF5A.

However, many questions remain opened, among these how this particular residue is synthesized in this archaeal strain, since the modification pathway in Sulfolobaceae still remains obscure. It is important to remember indeed that no homologue of the second enzyme involved in hypusination in eukaryotes, deoxyhypusine hydroxylase (DOHH), was found in any archaeal genome or proteome to date. For this reason, the attention focused on the characterization of the components of the hypusination pathway, the aIF5A protein and the homologue of the first enzyme of modification deoxyhypusine synthase (DHS), from structural and functional point of view.

3.2 Purification of recombinant aIF5A in *E. coli* and in *S. solfataricus*

In order to study these proteins, it is important first of all to obtain the recombinant proteins. Starting with aIF5A, in a previous work (Bassani *et al*, 2018) the gene coding for aIF5A (ORF Sso0970) had been cloned into two different expression vectors to allow the production of the protein in *E. coli* and in Sso. The pETM11 vector was used to introduce a 6x His-tag at the N-terminal of the protein in *E. coli* ROSETTA (DE3)/pLysS strain, whereas the pMJ05-ptf55 plasmid was used to introduce 6x His-tag at the C-terminal of aIF5A directly in *S. solfataricus* PH1-16 (ΔpyrEF). Both recombinant proteins were purified as described in Materials and methods.

As expected, the yield of recombinant protein produced in *E. coli* is rather high and around 20 mg of pure N-His-aIF5A were obtained from a 2L culture. However, this protein is in an unmodified version, being *E. coli* unable to perform the post translational modifications. On the other hand, the recombinant aIF5A-C-His produced in Sso was obtained in a much lower amount but was fully hypusinated (Bassani *et al*, 2018). In this work, both versions of the proteins, the unmodified and the hypusinated, were purified, and used in functional and structural studies. As shown in the figure 18A, the specific antibody anti-aIF5A is able to recognize the N-His-aIF5A and the aIF5A-C-His proteins, as well as the native aIF5A in a *Sulfolobus solfataricus* lysate. The native protein aIF5A has a molecular weight of 14491.8 Da, while the protein cloned in pETM11 and expressed in *E. coli* has a molecular weight of 17761.37 Da, as at the N-terminal of the protein there is the His-tag followed by a peptide specifically recognized by the TEV protease. The aIF5A-C-His expressed in Sso is purified as a protein of theoretical mass of 15472.85 Da. Figure 18B shows the SDS-PAGE analysis followed by Comassie Blue Staining of both Sso-aIF5A His-tagged proteins. Results show that the N-His-aIF5A purified from *E. coli* is a pure protein, while the aIF5A-C-His from *S. solfataricus* always co-purifies with another protein whose MW is around 35 kDa. Both aIF5A proteins were subsequently analyzed by Western Blot, using two antibodies: the specific anti-aIF5A (Fig. 18C) that recognizes both recombinat proteins, and the anti-hypusine (Fig. 18D) that recognizes only the protein produced in Sso, confirming the hypusination of aIF5A-C-His in Sso. To validate this finding and to clarify the nature of the modification, the mass spectrum of intact aIF5A-C-His from Sso PH1-16 (pMJ05-aIF5A-C-

His) was analyzed. The MS spectrum (Fig. 18E) revealed a main peak, relative to the most abundant polypeptide, with a mass of 15559.97. When compared with the theoretical mass of aIF5A-C-His, the mass increase of 86.77 correlated with the molecular weight of the hypusine residue.

Hence, we concluded that most of the recombinant aIF5A-C-His produced in Sso PH1-16 is hypusinated.

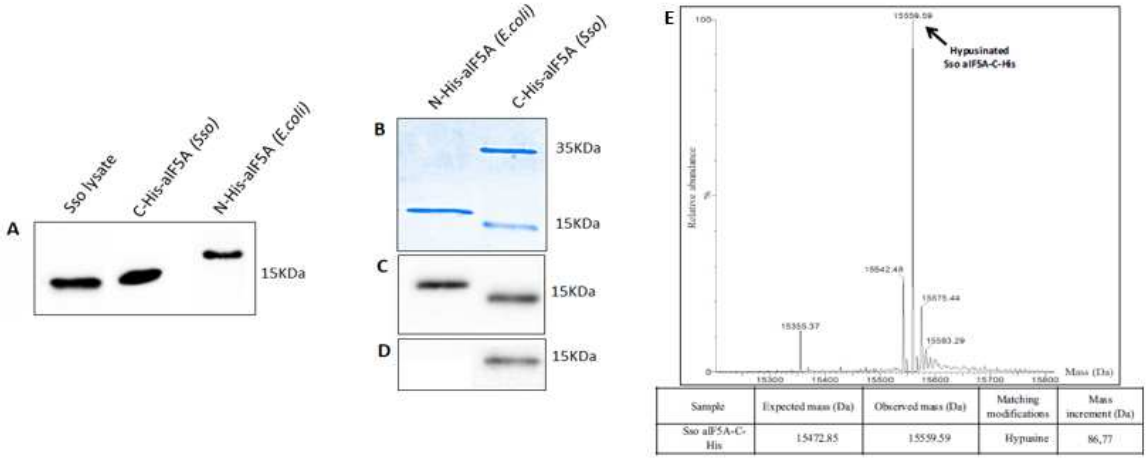


Fig. 18 Detection of recombinants aIF5A proteins. (A) Western Blot analysis of native protein form Sso lysate, aIF5A-C-His from Sso and N-His-aIF5A from *E. coli* with anti-aIF5A. Recombinant proteins purified from Sso and *E. coli* and detected with Comassie Blue Staining (B), Western Blot using anti-aIF5A (C), and with anti-hypusine (Merk-Millipore) (D). (E) MS spectrum of recombinant aIF5AC-His produced in Sso PH1-16 (pMJ05-aIF5A-C-His) (Bassani *et al*, 2018).

3.3 Production and purification of *Sulfolobus solfataricus* deoxyhypusine synthase (DHS) in *E. coli*

The presence of the hypusine modification of aIF5A suggests a similar modification pathway between Archaea and Eukarya. In order to elucidate the *Sulfolobus solfataricus* post-translational modification mechanism, the characterization of the deoxyhypusine synthase enzyme (Sso DHS) was undertaken.

The Sso DHS gene (ORF Sso0967) was cloned and overexpressed in *E. coli*, to obtain a proper amount of recombinant protein to use for structural and functional studies.

The ORF Sso0967 was inserted into two different expression plasmids (pETM11 and pQE-70), in order to obtain DHS enzyme His-tagged either at the N-terminal or at the C-terminal

position respectively. First, analyzing the expression levels of both recombinant proteins, a different solubility degree was noticed, due to the presence of His-Tag at N-terminal or C-terminal position. When lysates, overproducing the two version of the protein, were subjected to centrifugation, insoluble proteins localized in pellet while soluble proteins are situated in supernatant. Both fractions were separated by SDS-PAGE, together with the eluate of the Ni-NTA affinity purification. Figure 19 shows that the N-His-DHS is mainly produced in insoluble form, localizing in the pellet (Fig.19, lane 1), whereas the DHS-C-His is more soluble (Fig.19, lane 6). The molecular mass of the DHS in denaturant elecrophoresis is 35 kDa. On the basis of this result only the C-terminal His-tagged protein was used for following experiments.

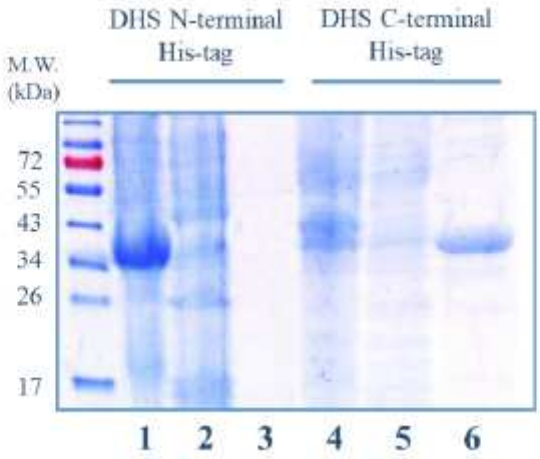


Figure 19. Expression and purification of Sso DHS in E. coli. Lanes 1-3: DHS N-terminal His tagged in *E. coli* ROSETTA (DE3)/ pLysS; lane 1: total proteins in the pellet; lane 2: total proteins in the supernatant; lane 3: N-His-DHS eluate after affinity purification. Lanes 4-6: DHS C-terminal His tagged in *E. coli* BL21 (DE3); lane 4: total proteins in the pellet; lane 5: total proteins in the supernatant; lane 6: DHS eluate after affinity purification

3.4 Structural characterization of aIF5A and DHS

With the aim to achieve structural characterization of aIF5A and DHS proteins from *S. solfataricus*, two techniques have been employed: Small-angle X-ray scattering (SAXS), that has emerged as an important method for studying the structural properties of proteins in solution, also at high temperature, and size-exclusion chromatography.

3.4.1 N-His-aIF5A is a monomer in solution and is indipendent of the hypusine residue

Small-angle X-ray scattering (SAXS) is a technique suitable for the analysis of the overall shapes, dimensions and oligomeric states of macromolecules in solution. Therefore, SAXS measurements were used to probe the molecular shape of N-His-aIF5A (from *E. coli*) directly in solution and exploiting the possibility to carry out measurements at high temperature in order to obtain structural details about this highly thermophilic protein in a "close to physiological" state. As described in Materials and Methods, all SAXS curves were recorded at Diamond Light Source (Didcot, UK), using the unmodified protein expressed in *E. coli,* with a 6 x Histidine flag at the N-terminal position. It was used only this recombinant protein because for this type of analysis a large quantity of protein is required (in total 1.5 mg of protein), an amount impossible to obtain from that expressed in *Sulfolobus*, since as told above, the yield of purification of aIF5A-C-His is low.

The N-His-aIF5A was analyzed at different conditions: with 60mM and 1M of KCl, at five different concentrations (0.5, 1, 2, 5, 10 mg/ml) and at nine different temperatures (20, 25, 30, 35, 40, 45, 60, 63°C), to investigate changes in the oligomeric form or state of the protein in function of different salt concentrations and at different temperatures and to see the stability of the *S. solfataricus* proteins, being a thermophilic organism, thus suitable for living under extreme conditions of life. 63°C was the maximum temperature achieved by the instrument, and it was the most interesting to study, as very close to the normal growth temperature of this archaeal organism. The X-ray scattering curve obtained for N-His-aIF5A is shown in figure 20.

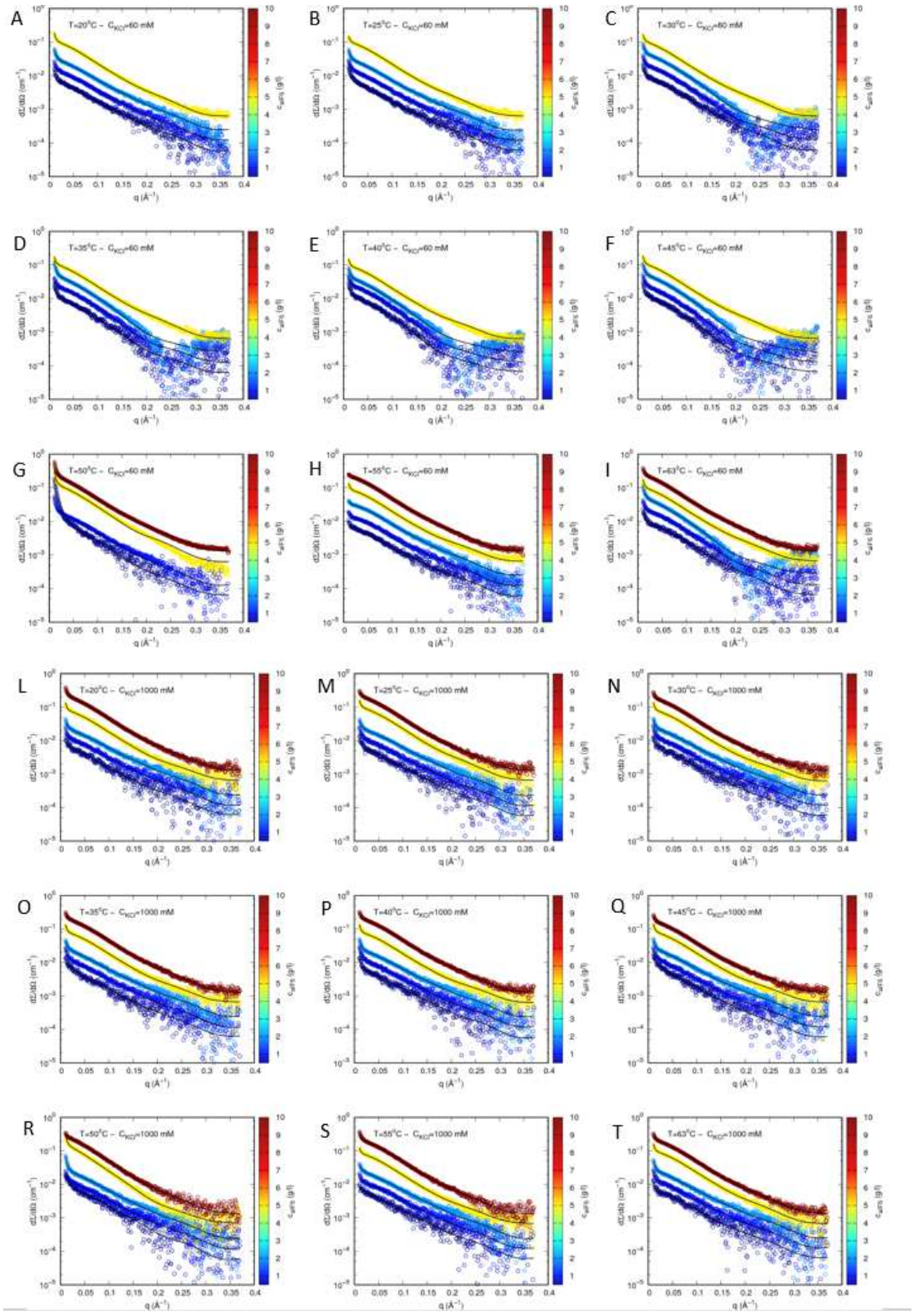


Fig.20 SAXS curves of N-His- aIF5A at 60 mM KCl (A → I) and at 1M KCl (L → T).

SAXS profiles are represented at increasing concentrations of the protein (highlighted by the various colors), keeping the concentration of KCl (60mM in the panels from A to I and 1M KCl in the panels from I to T) and temperature (20°C (panels A and L), 25°C (panels B and M), 30°C (panels C and N), 35°C (panels D and O), 40°C (panels E and P), 45°C (panels F and Q) , 50°C (panels G and R), 55°C (panels H and S), 63°C (panels I and T).

It can be noted that measurements performed at low concentrations of the aIF5A protein (0.5, 1, and 2 mg/ml) show a noisy signal, therefore difficult to interpret. Measurements performed at high protein concentrations (5 and 10 mg/ml) provide a marked signal, due to the presence of a higher number of particles in solution.

Furthermore, all the graphs show a high signal at small values of the scattering vector q (at the left part of each pictures) that indicates the probable presence of large aggregates of the proteins in solutions.

To verify this initial evaluation, it is necessary to perform a data analysis using a model that combines the Guiner's law with Porod (fit Guinier-Porod). From this analysis two important informations on the structure of the protein are obtained: the radius of gyration and the molecular weight of aIF5A. The radius of gyration Rg of a protein is defined as the root mean square distance from each atom of the protein to their centroid so it is an indication of the dimension of the protein. All the curves have been considered and from this analysis the radius of gyration of the protein has an average of 22.4 Å (2.24 nm), consistent with the monomeric form of the N-His-aIF5A (Dias *et al*, 2013).

This monomeric conformation was also confirmed with the molecular weight analysis that in all cases is in a range of 15000-17000 Da, as the molecular mass predicted from Expasy tool starting from the primary structure of the protein. These observations were found in all samples with different conditions, and demostrate that the protein does not change the oligomeric state in presence of high salt concentrations or at low and high temperatures.

The Kratky plot, I(q)*q2 vs. q plot, is thus informative to check globularity and flexibility of proteins. In the case of well-folded globular protein, the Kratky plot will exhibit a "bell-shape" peak at low q and converges to the q axis at high q (figure 21A). The Kratky plot will not converge to the q axis if the protein has a pronounced flexibility. Proteins with multiple-domains could have additional peak(s) (or shoulders) at low q but the Kratky plot will still converge to the q axis at high q.

For this analysis only the curves at 63°C are considered, at each protein concentration, and at both 60 mM KCl (Fig.21B) and 1 M KCl (Fig.21C).

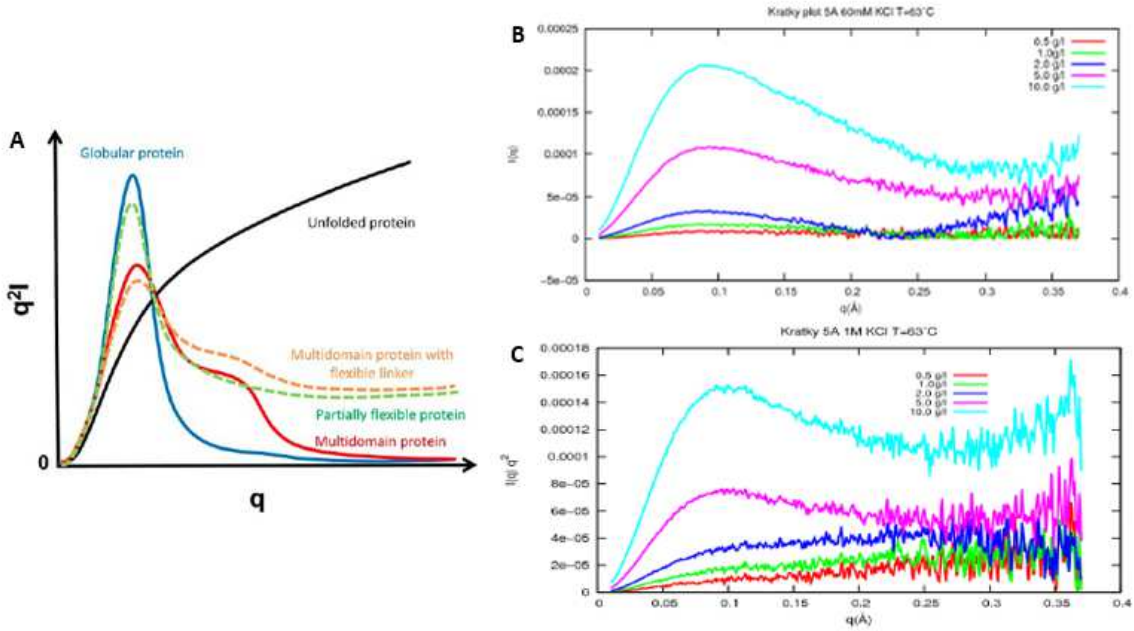


Fig.21 Kratky plots. (A) Schematic representation of typical Kratky plots. Curvature is depending on molecular shape, degree of flexibility. (B) Kratky plots of Sso N-His-aIF5A at increasing concentration, 60 mM KCl, at 63°C. (C) Kratky plots of Sso N-His-aIF5A at increasing concentration, 1 M KCl, at 63°C.

As shown in the graphs at high concentration of protein (5, 10 mg/ml) we can better appreciate the trend of the curve, which, in both concentrations of KCl, corresponds to a partially denatured state of the aIF5A. This means that the protein could be composed of rigid and other more flexible domains.

The tertiary structure of *Sulfolobus solfataricus* aIF5A was determined using the QUAFIT software, which analyzes the SAXS data considering the hypothesis that the protein consists of rigid domains, whose crystallographic structure is known and of flexible domains. QUAFIT realizes a SAXS curves fitting, optimizing the relative position of the rigid domains and the angles of Ramachandran relative to the flexible domains. It is possible to analyze the SAXS curve assuming that the protein can be distributed over multiple structures, each having a distinct relative position of the rigid domains and distinct conformations of the flexible domains.

Below it is shown the amino acid sequence of the recombinant aIF5A expressed in *E. coli*, used for SAXS measurements: the primary structure was obtained from the UniProt database (Q97ZE8), and at the N-terminal position are shown underlined the extra amino acids result from the pETM11 vector, in red the six histidines tag and in blu the recognition site for Tobacco Etch Virus (TEV) protease.

ML**HHHHHH**PMSDYDIPTT**ENLYFEGA**MSITYTTVGELKVGSYVVIDGEPCRVVE
VTKAKTGKHGSAKANVVAIGVFSGAKKTLMAPVDQQVEVPIIEKHIGQIIADMGN
KIQVMDLESYETFEIEKPTEDELASKIKPNAELEYWEIMGRRKIVRVK

The N-His-aIF5A is thus composed of 157 amino acids.

To determine the protein structure with QUAFIT to identify rigid domains and flexible linkers is necessary. The following domains was defined, through homology sequence with analogous proteins (Yao *et al*, 2003):

• first rigid domain: residues 27-93;

• second rigid domain: residues 98-154;

• first flexible linker: residues 1-26;

• second flexible linker: residues 94-97;

• third flexible linker: residues 155-157

The QUAFIT analysis was performed on the curve of N-His-aIF5A 10 mg/ml, at 63°C and with 60 mM KCl, as is the closest to the physiological condition of the protein in *Sulfolobus.* Ten probable models of the tertiary structure of the aIF5A protein were produced by QUAFIT analysis. A representation of these 10 conformers is shown in Figure 22: in red is represented the N-terminal domain, in green the C-terminal domain, while in blue the extra amino acids of the protein (6 x His-tag, TEV site).

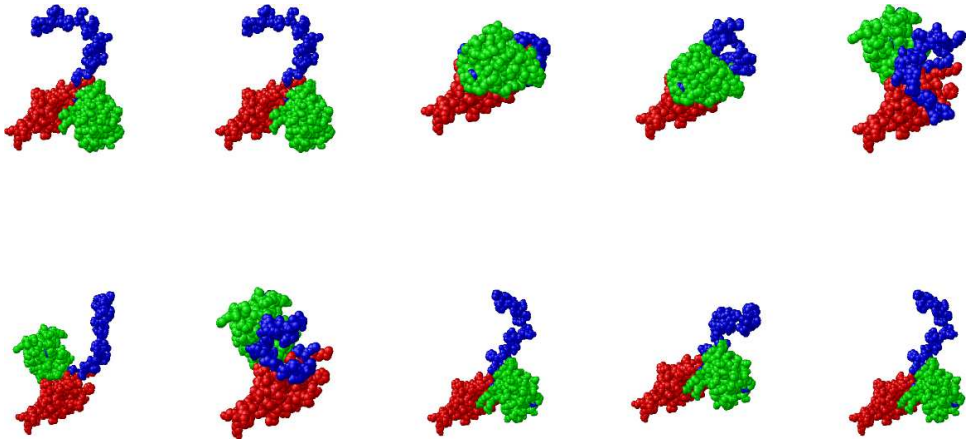


Fig 22 Ten possible conformers of Sso N-His-aIF5A achieved with QUAFIT software (Spinozzi and Beltramini, 2012).

Unfortunely this latter blue part (about 10% of the total protein) interferes notably with this type of analysis, so a suitable fit with other homologous structures available from X-ray data on PDB was not possible.

Concluding, with SAXS analysis many structural informations were obtained: aIF5A protein is a monomer in solution, probable made up by a cluster of 10 possible conformations in which more rigid domains (N and C terminal ends) are connected by more flexible parts (for example linkers and loops), in agreement with the crystal structure of the a/eIF5A homologue proteins in Eukarya and other Archaea. Furthermore, no difference was found on the overall shape of the protein changing the conditions for protein solubility, such as temperature and salt concentration. Thus, aIF5A is a stable monomeric termophilic protein, however the removal of the histidine tail would be required for a better SAXS analysis of the protein.

In yeast it was demostrated that eIF5A exists as a dimer in solution, dependently on RNA and probably on hypusine residue (Gentz *et al*, 2009; Dias *et al*, 2013). The monomeric conformation of N-His-aIF5A protein has just been shown by SAXS data. Nevertheless, this recombinant protein is expressed in *E. coli* and certainly does not harbor the post translational modification. Moreover N-His-aIF5A does not hold *S. solfataricus* RNA molecules that could interact with the protein during the purification and enable the oligomerization of the protein, as demostrated in *S. cerevisiae*. Given the structural similarity between the archaeal and the eukaryal protein, to investigate this hypotesis in *Sulfolobus*

solfataricus, the C-terminal His-tagged aIF5A, affinity purified from Sso PH1-16, was subjected to size-exclusion chromatography in a minimal buffer containing 50 mM Tris HCl pH 7.4 and 150 mM KCl, using a Superdex 75 column of 24 ml. It was loaded on the column also the N-His-aIF5A (produced in *E. coli*) at the same conditions. Bovine serum albumin (BSA) and myoglobin were used as molecular mass standards to calibrate the column. BSA (67 kDa) and myoglobin (17 kDa) eluted in 1.9 ml and 13.3 ml respectively. Figure 23 shows that N-His-aIF5A (not containing the hypusine residue) eluted at 13.3 ml (Fig.23A) and aIF5A-C-His (containing the hypusine residue) eluted at 13.33 ml (Fig.23B), both corresponding to a 17 kDa protein, below the myoglobin peak, and is approximately the size expected for the aIF5A monomer. Western Blot analysis using anti-aIF5A antibodies (Fig. 23C) of the chromatographic protein fractions confirmed that both unmodified and hypusinated proteins were eluted in a column volume corresponding to a 17 kDa protein. These findings are consistent with the result observed with SAXS and confirm that Sso aIF5A is a monomer and is indipendent of the hypusine residue.

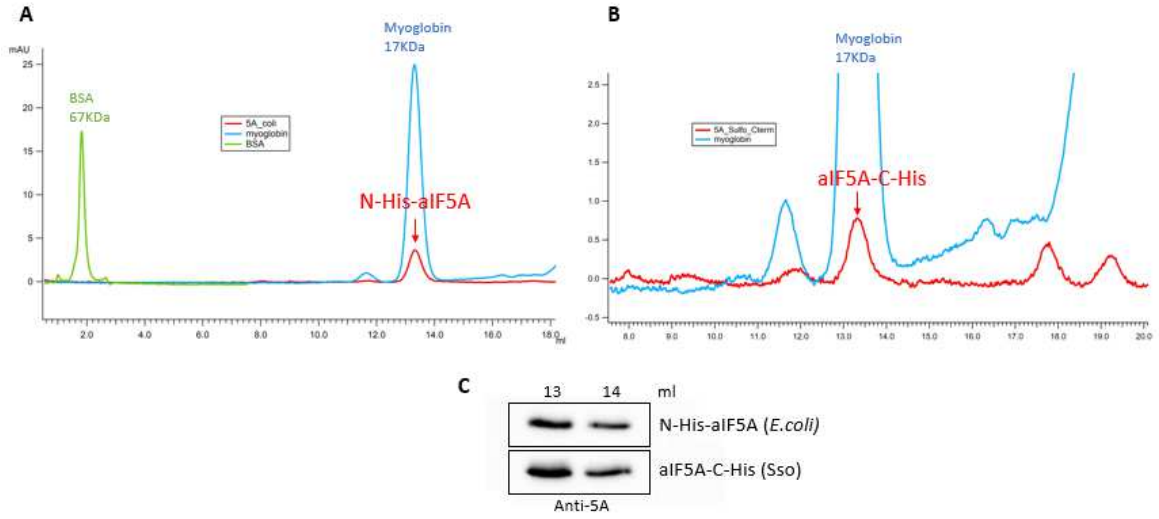


Fig.23 Size exclusion chromatography on Superose 75. (A) Size-exclusion chromatography of N-His-aIF5A (expressed in *E. coli*), BSA and myoglobin as standars. (B) Size-exclusion chromatography of aIF5A-C-His (expressed in Sso) and myoglobin as standard. (C) Western Blotting of chromatographic protein fractions with anti-aIF5A antibody.

3.4.2 DHS-C-His is a tetramer as the eukaryotic DHS

Eukaryotes DHS exists as a tetramer of four identical subunits of 40 kDa to 43 kDa, depending on the species. So far human DHS has been better characterized. Crystal structure shows that the DHS tetramer contains two pairs of tightly associated dimers, A1A2 and B1B2 with two active sites, located in each dimer interface. The N-terminal tail from one monomer interacts with the N-terminal of the other monomer creating a movable channel for the entry of spermidine and IF5A precursor to the active-site cavity (Liao *et al*, 1998). Hypothesizing a structural similarity between the eukaryotic and the archaeal enzyme, these findings could explain why a solubility problem was found in the expression and purification of *Sulfolobus solfataricus* DHS in *E. coli* only with the flag of six histidine at the N-terminal position of the recombinant protein. Probably this additional tag at the N-terminal interferes with the proper folding of the Sso enzyme. An important achievement was obtained from the characterization of an archaeal DHS enzyme in *Haloferax volcanii* as a tetramer of 148 kDa, that underlines even more the structural closeness between the archaeal and eukaryotic enzymes. As nothing is known about the *Sulfolobus* DHS enzyme, in order to investigate the structural features, a SAXS analysis of Sso DHS-C-His was performed at 63°C. Figure 24 shows the SAXS experimental data and the computed scattering curves based on the crystallographic structures of human DHS (Fig. 24B), considering all possible conformation (monomer, dimer, trimer and tetramer hDHS). The scattering curve computed based on the crystallographic structure of human DHS tetramer exhibited the best fit to the experimental data (Fig. 24C), compared to the other scattering curves (data not shown).

This result proved for the first time that DHS enzyme from *S. solfataricus* is a tetramer, like the eukaryotic counterpart, and probably the N-terminal tail of the enzyme plays a key role in the correct folding of the protein, to acquire the active conformation of DHS.

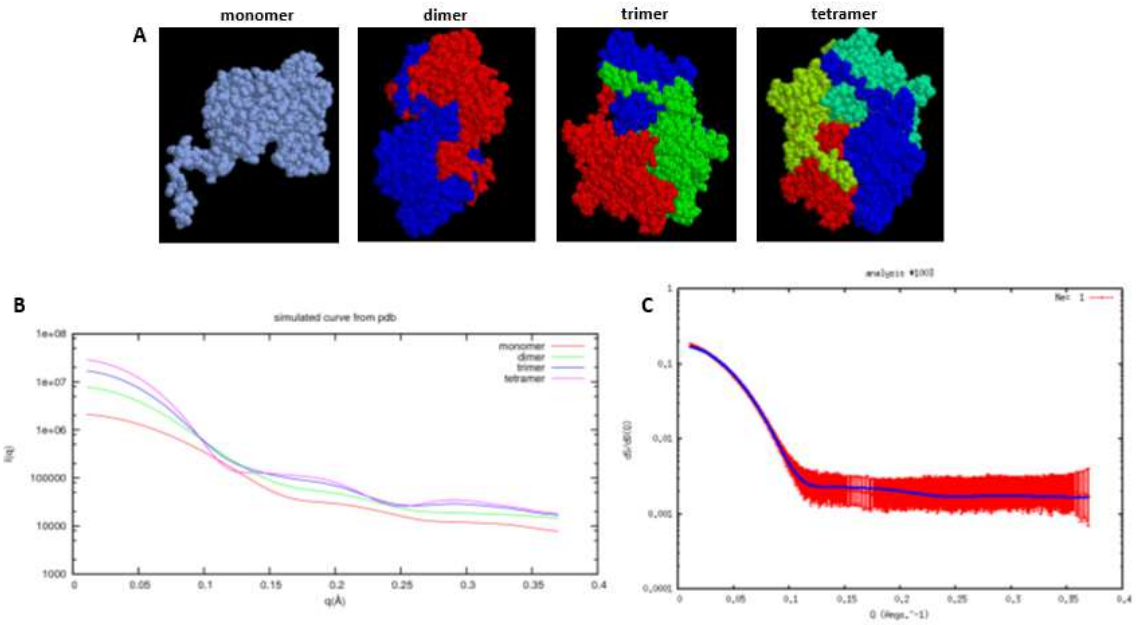


Fig.24 Small-angle X-ray scattering (SAXS) curves of DHS enzyme. (A) Rapresentations of the crystal structure of human DHS, considering the monomer, dimer, trimer and tetramer conformations (Liao *et al,* 1997). (B) Computed scattering curves based on the crystallographic structures of human DHS. (C) Experimental scattering curve of Sso DHS-C-His superimposed on the computer scattering curve based on crystallographic structure of human DHS tetramer.

3.5 Complex formation between aIF5A and DHS

We demonstrated that aIF5A is hypusinated and a monomer in solution, while DHS is a tetramer, as it was shown already for the eukaryotic DHS enzyme.

Next step was to investigate if these two recombinant proteins interact *in vitro* to form a complex. To demonstrate it, three different techniques were used: SAXS analysis, Size-exclusion molecular chromatography and native PAGE. The recombinant Sso proteins N-His-aIF5A and DHS-C-His both purified from *E. coli,* were incubated together at 65°C for 1 hour, to allow their interaction. SAXS measurements were useful technique not just to analyze structural features of the proteins in solution, but also to detect the complex formation among two different proteins. The analysis was carried out at 63°C as described above, at Diamond Light Source syncrotron (Didcot, UK) and figure 25 shows the preliminary analysis of the DHS and DHS-aIF5A curves. A shift of the experimental curves between DHS alone (violet line), and DHS with 5A (blue line), is visible, in particular in the first part of the graps at low q, that is an indication of a change in the molecular weight of proteins, a strong clue that the aIF5A protein was binding to DHS enzyme. The number of

active sites for aIF5A-binding per DHS-tetramer has not been characterized biochemically, so actually we do not know how many aIF5A molecules bind each DHS protein.

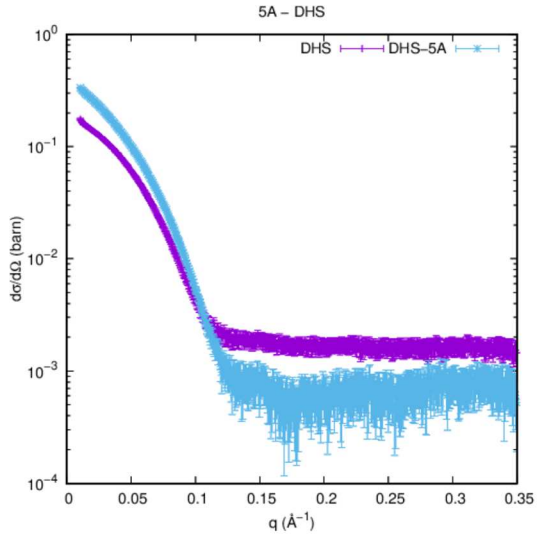


Fig. 25 Small-angle-X-ray scattering (SAXS) curves of DHS-C-His (from E. coli) alone (violet line) and DHS-C-His pre-incubated together with N-His-aIF5A (from E. coli) at 65°C for 1h (blue line).

To confirm the complex formation, a size-exclusion chromatography was performed. DHS-C-His and N-His-aIF5A expressed in *E. coli* were loaded first alone, to detect their characteristic elution volume, and then loaded together, pre-incubated at 65°C for 30 min, with all other components for the reactions (spermidine and NAD^+), into 3 ml Superdex 200 increase 5/150 GL Tricorn column, which was set to work at high temperature, in particular at 65°C. This feature is important for experiments with DHS-C-His protein because, unlike aIF5A, it is a more instable protein and has a tendency to precipitate at low temperature condition.

Figure 26A shows the chromatogram obtained from different chromatographic runs of aIF5A alone (pink), DHS alone (blue), DHS+NAD+Spermidine (green) and DHS+aIF5A+NAD+Spermidine (red). All reactions were performed at the same conditions and were overlapped to show differences in elution profiles. It is clear that the presence of both proteins preincubated together, shift the peak of DHS protein (in red) at high molecular weight positions (around 140 kDa, confirming once again the tetrameric form of Sso DHS), and the peak corresponding to aIF5A decrease almost all, probably becouse most of the aIF5A protein was linked to the enzyme. Anyhow, aIF5A is a very small protein (17 kDa)

and a change in a molecular weight is very difficult to predict also in this case. Furthermore, the size-exclusion column used for this analysis is very small (3 ml), that is unfit to discriminate small changes in molecular mass of proteins.

However, SDS-PAGE followed by Western Blot analysis of the fractions of each run was performed to confirm the complex formation between aIF5A and DHS. As shown in figure 26B, an anti-Histidine antibody, that recognize the tag of six histidine present in each protein, was used to detect either N-His-aIF5A and DHS-C-His. At lower Western Blot panel (in pink), N-His-aIF5A elutes in a volume corresponding to 17 kDa protein, approximately the size expected for the aIF5A monomer, while in the highest panel (in blue) DHS-C-His elutes in a volume corresponding to the tetrameric form of DHS. In the middle red panel both proteins are present in the gel at the same position, Western Blot shows that both DHS-C-His and N-His-aIF5A elute together in a column volume corresponding to tetrameric form of DHS and aIF5A was not found anymore in 17 kDa position, but becomes DHS-bound.

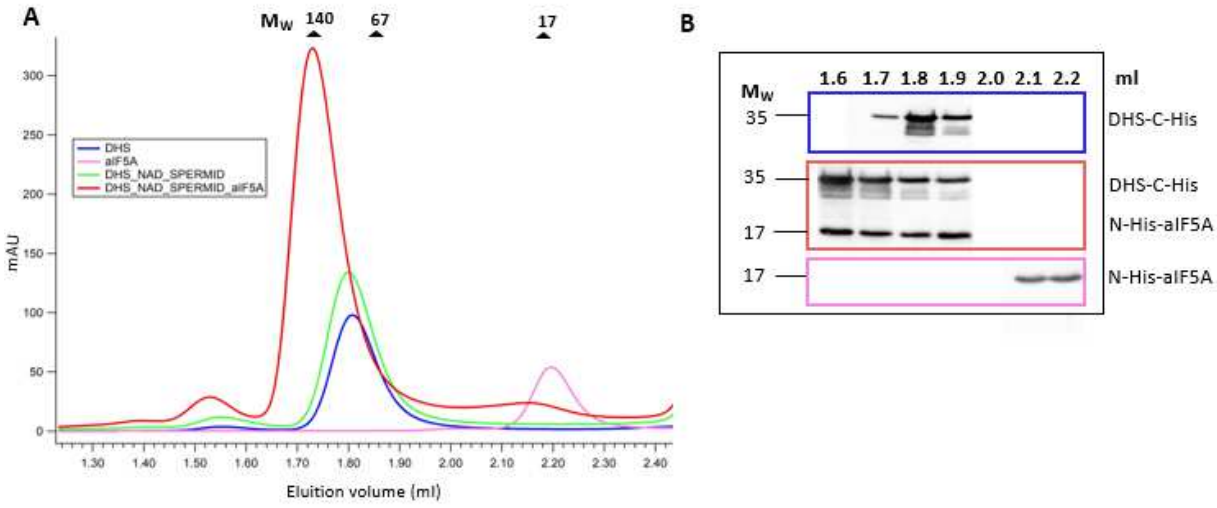


Fig. 26 Complex formation between N-His-aIF5A and DHS-C-His with size-exclusion chromatography. (A) Gel filtration profiles of 20 µg aIF5A alone (in pink), 20 µg DHS alone (in blue), 25 µg DHS+2 mM NAD + 2 mM Spermidine (in green) and 50 µg DHS + 25 µg aIF5A + 2 mM NAD + 2 mM Spermidine (in red) on Superdex 200 increase 5/150 GL Tricorn column. Myoglobin (17 kDa), BSA (67 kDa) and lactate-dehydrogenase (140 kDa) were used to calibrate the size-exclusion elutionfractions (▲). (B) SDS-PAGE/Western Blot analysis with anti-histidine antibody of all the chromatographic fractions beside each chromatogramm.

Definitive evidence for the formation of a stable complex of the enzyme DHS and the substrate aIF5A, was obtained from non-denaturing gel electrophoresis (Figure 27A). Since the theoretical pI of the enzyme DHS (6.9) is close to that of aIF5A (6.15), molecular size should be the major factor determining the electrophoresis mobility of these proteins in the native state. As expected, the tetrameric enzyme (denoted E) was separated from the much smaller protein substrate (denoted S) upon native gel elecrophoresis (Figure 27A, lane 1-2). In a mixture containing the DHS-C-His enzyme (E) and N-His-aIF5A (S) at a ratio of 1:4 (molar rario) in 200 mM glycine/ NaOH pH 8.2, a new band of protein (denoted E•S), migrating a little more slowly than the enzyme alone, was observed and at the same time the band corresponding to the aIF5A protein shows in figure 27A, lane1 seems to disappear in Figure 27A, lane 3-4. The formation of E-S complex is not dependent on spermidine and NAD^+, suggesting that binding of these two cofactors to the enzyme is not a prerequisite for its association with the protein substrate. The identity of this band as a complex was determined by elution and electrophoresis under denaturing conditions, as shown in figure 27B. The proteins from E•S band in figure 27A (lane 4), highlighted by a red square, dissociated into the DHS monomer (35 kDa) and aIF5A (17 kDa) upon SDS-PAGE, providing direct proof that these bands contain a complex of enzyme and protein substrate (Figure 27B, lane 3).

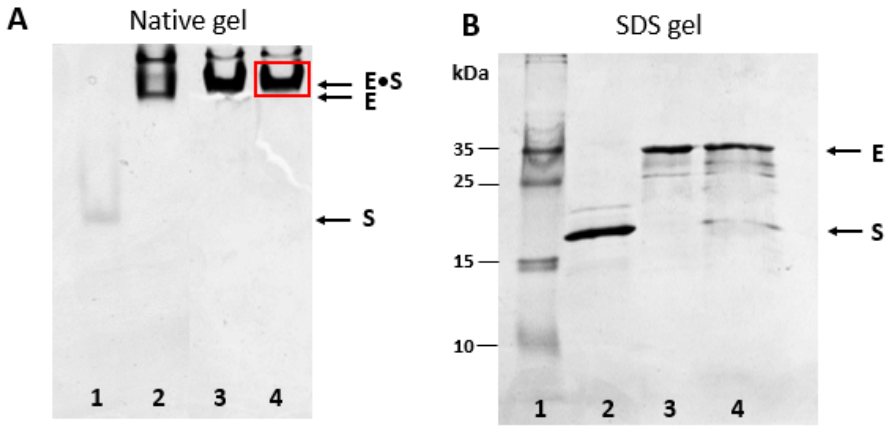
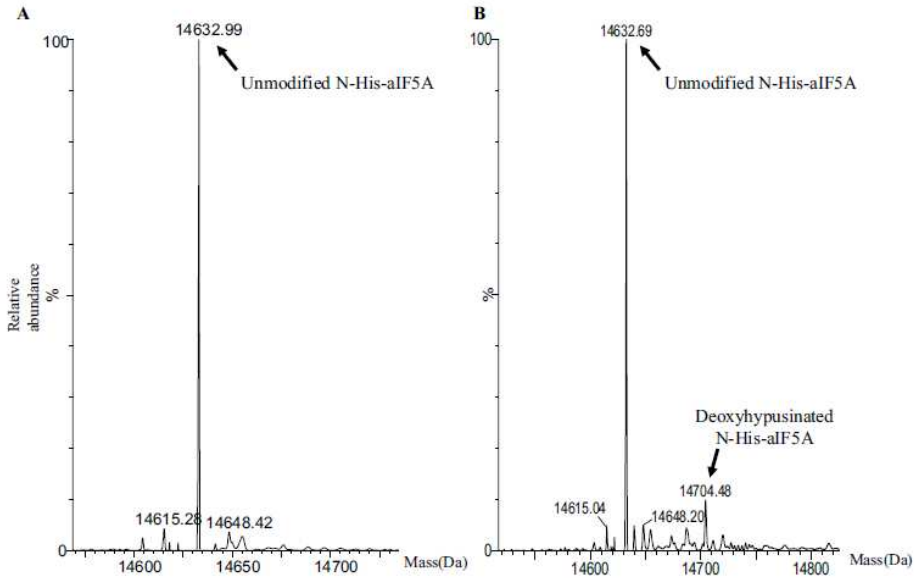


Fig.27 Detection of the DHS/aIF5A complex formation by (A) non-denaturing PAGE and (B) SDS-PAGE (A) Native PAGE electrophoresis on 12% native gel. Lane 1: N-His-aIF5A alone; lane 2: DHS-C-His alone, lane 3: N-His-aIF5A plus DHS-C-His; lane 4: N-His-aIF5A plus DHS-C-His and 1mM NAD^+ and 1mM spermidine (complete reaction). (B) SDS-PAGE on 15% SDS-gel. Lane 1: molecular mass standard; lane 2: N-His-aIF5A alone; lane 3: DHS-C-His alone; lane 4 complex (E•S) band excised from the gel shown in A, lane 4.

3.6 Sso DHS performs the deoxyhypusine sythesis *in vitro*

Results described in the previous section, prove that the *S. solfataricus* proteins DHS and aIF5A interact *in vitro* to form a stable complex, suggesting that DHS enzyme is, in all probability, involved in the modification of aIF5A.

The absence of a DOHH enzyme homologue in Archaea prompted us to investigate whether DHS is responsible for both reactions as the enzyme discovered in *Trichomonas vaginalis* (Quintas-Granados *et al,* 2015). To confirm this finding, the Sso DHS activity was tested with an *in vitro* modification assay in which all components of the reaction, including the two recombinant proteins from *E. coli*, N-His-aIF5A and DHS-C-His and cofactors NAD^+ and spermidine, are incubated together for 1 h at 65°C, and analyzed by mass spectrometry. Figure 28 shows the spectra of N-His-aIF5A after incubation with (Fig. 28A) and without (Fig. 28B) the presence of DHS enzyme. An additional peak corresponding to the deoxyhypusinated form of N-His-aIF5A was observed in presence of DHS-C-His (Fig.28B). Even if the amount of intermediate is low (around 10% of total aIF5A), an important information is obtained, that the DHS enzyme is involved in the modification of aIF5A in *S. solfataricus*, catalyzing the formation of the intermediate (deoxyhypusine-aIF5A) of the reaction, excluding the hypotesis of bifunctional activity of this enzyme. So, we can conclude from this experiments that the first step of hypusine modification is conserved and very similar to the eukaryotes.



Sample	Expected mass (Da)	Observed masses (Da)	Matching modifications	Mass increment (Da)
Unmodified N-His-aIF5A	14632.83	14632.99		+16
N-His-aIF5A in the presence of aDHS-C-His	Unmodified N-His-aIF5A (14632.83)	14632.69		−14
	Deoxyhypusinated N-His-aIF5A (14703.83)	14704.56	Deoxyhypusine	+71
	Hypusinated N-His-aIF5A (14719.83)			

Fig.28 The unmodified Sso-5A is in vitro deoxyhypusinated by the recombinant Sso-DHS enzyme.
(A) MS control spectrum of the unmodified recombinant Sso-5A; the protein was incubated with all components of the in vitro hypusination assay, without the DHS enzyme. (B) MS spectrum of the recombinant Sso-5A subjected to the in vitro hypusination assay.

3.7 *Sulfolobus solfataricus* aIF5A is a multifunctional protein

3.7.1 aIF5A is involved in translation process

Among the small group of universally conserved translation factors, the eukaryotic eIF5A and its bacterial orthologue EF-P have been shown to stimulate translation elongation of proteins with proline stretches, rescuing stalled ribosomes (Ude *et al*, 2013; Schuller *et al,* 2017). So far, an involvment of archaeal aIF5A in protein synthesis has not been demostrated yet. To investigate if *S. solfataricus* aIF5A, consistent with his name of translation factor, partecipates in protein synthesis, different tests were carried out.

We first performed the immunolocalization of native aIF5A during different steps of lysate fractionation: when a Sso lysate (called S30) is subjected to an ultracentrifugation (100000 g), the resultant supernatant contains the cytoplasmatic portion (called S100), whereas the pellet contains crude 70S ribosome, with proteins strongly bound to 70S still attached. If these ribosomes are subjected to high (500 mM) salt wash and centrifuged again, the supernatant contains all those proteins dissociated from the ribosomes (called HSW, High Salt Wash), and the pellet contains washed 70S. All these fractions were separated on SDS-PAGE and the presence of aIF5A was detected using anti-aIF5A and anti-hypusine antibodies, as shown in figure 29:

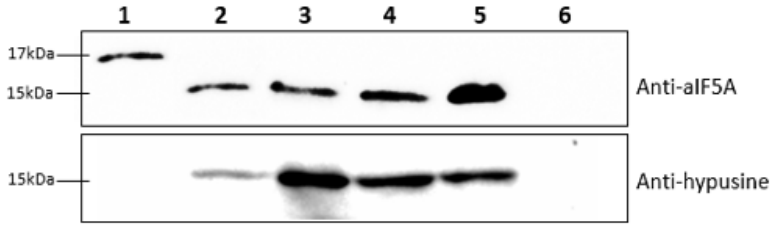


Fig.29 Cellular localization of the native Sulfolobus solfataricus aIF5A Proteins were separated on 15% SDS-gel and aIF5A was detected with Western Blot using anti-aIF5A and anti-hypusine antibodies. Lane 1: N-His-aIF5A 10 pmol as a control for the antibodies, lane 2: 20 µg S30 lysate, lane 3: 20 µg S100 fraction, lane 4: 50 µg HSW (high salt wash) fraction, lane 5: 50 pmol of crude ribosomes 70S, lane 6: 50 pmol washed ribosomes 70S.

The protein appears to be present either in post ribosomal fractions (Fig.29, lane 3-4) but also in crude 70S ribosome (Fig.29, lane 5), without significant differences between the two antibodies while it is absent from ribosomes which have been high salt washed (Fig.29, lane 6). Recombinant N-His-aIF5A was loaded as a control (Fig.29, lane 1).

The localization of native aIF5A in 70S fractions is a first indication of a possible interaction with ribosomes. To better investigate this hypotesis, the washed 70S ribosomes, free of detectable native aIF5A, were dissociated in the two subunits and 30S/50S were isolated separately using a sucrose gradient. These *S. solfataricus* purified ribosomal subunits were used in an *in vitro* assay to test a possible binding of recombinant aIF5A to 30S and 50S separately. Although the homologous proteins act as translation elongation factors and attend its function only once the active 70S ribosomes are formed, we investigate if aIF5A has a possible binding preference for one of the two subunits as the very first analysis. It was

demonstrated indeed that the post-translational modification in eIF5A and EF-P play a key role for the function of the proteins as elongation factors (Schmidt *et al*, 2016; Bullwinkle *et al*, 2013), so either the unmodified version produced in *E. coli* and the hypusinated one produced directly in *S. solfataricus* were used in this analysis.

Increasing amounts of purified 50S and 30S subunits (20, 40, 80 pmol) were incubated for 10 min at 65°C in a binding buffer (as described in Materials and Methods) with a fixed amount of both recombinant aIF5A proteins (30 pmol), to explore if the hypusine residue could mediate the interactions with ribosomes, as the eukaryal counterpart.

Resulting ribosome subunits / recombinant aIF5A complexes were separated from the unbound protein by ultracentrifugation and the supernatant fractions (unbound proteins) were withdrawn, loaded on SDS-gel and aIF5A in this fraction was detected by Western Blot (figure 30). Should aIF5A bind to the ribosomes, a decrease of free aIF5A detectable in the supernatant should be visible with increasing 30S and 50S concentrations.

Results are presented in figure 30. Using the unmodified protein expressed in *E. coli* (N-His-aIF5A) the amount of free aIF5A in supernatant is the same as the control without ribosomes (Fig.30, lane 1 and 5), either with 30S and 50S subunits, suggesting that this unhypusinated protein is unable to bind ribosomes (Fig.30, upper part). Using instead the hypusinated protein expressed in *S. solfataricus* (aIF5A-C-His) we observed, in the lower part of the figure 30, that with 30S ribosomes the results is the same just described, but interestingly when the aIF5A with hypusine residue was incubated with 50S subunits, probably a portion of aIF5A became 50S-bound, because we observe a decrease of aIF5A in supernatant compared to the control without 50S (Fig.30, lane 5), in particular at higher concentrations of 50S subunits (40 and 80 pmol) (Fig.30, lane 7-8).

Fig.30 In vitro interaction of recombinants aIF5A with S. solfataricus ribosomes isolated subunits. Lane 1: N-His-aIF5A without ribosomes, lane 2: N-His-aIF5A and 20 pmol 30S, lane 3: N-His-aIF5A and 40 pmol 30S, lane 4: N-His-aIF5A and 80 pmol 30S, lane 5: aIF5A-C-His without ribosomes, lane 6: aIF5A-C-His with 20 pmol 50S, lane 7: aIF5A-C-His with 40 pmol 50S, lane 8: aIF5A-C-His with 80 pmol 50S.

So, it is possible that aIF5A interacts with 50S ribosomes and likely the hypusine residue is important in this connection. But in order to clarify how this protein acts as a translation factor, further investigations are needed, in particular with *in vivo* assays that resemble the conditions of translation process.

For this reason, to analyze the behaviour and localization of native aIF5A during active translation, a *Sulfolobus solfataricus* lysate (S30) was programmed for protein synthesis. The programmed lysate was incubated at 70°C for 15 min, in the presence of exogenous mRNA and all other components necessary for translation and fixed with formaldehyde, to stabilize 70S ribosomes which are easily to dissociate in *S. solfataricus.* Then the lysate was fractionated on density sucrose gradients; the fractions were monitored for their absorbance at 260nm and probed by Western Blotting with anti-aIF5A. A control sample of a lysate not programmed for translation was analyzed as well.

The result in Figure 31A shows that in the control (not programmed for translation) the protein is exclusively present in the top low-molecular-weight fractions and active 70S peak is not detected. Interestingly, when ribosomes are engaged in active translation (Fig.31B), a portion become ribosome-bound, and aIF5A is found in a 30S, 50S and 70S ribosomal fractions. To detect the 50S fractions also aIF6 was probed, which is known to be stably bound to the large ribosomal subunits (Benelli *et al,* 2009b).

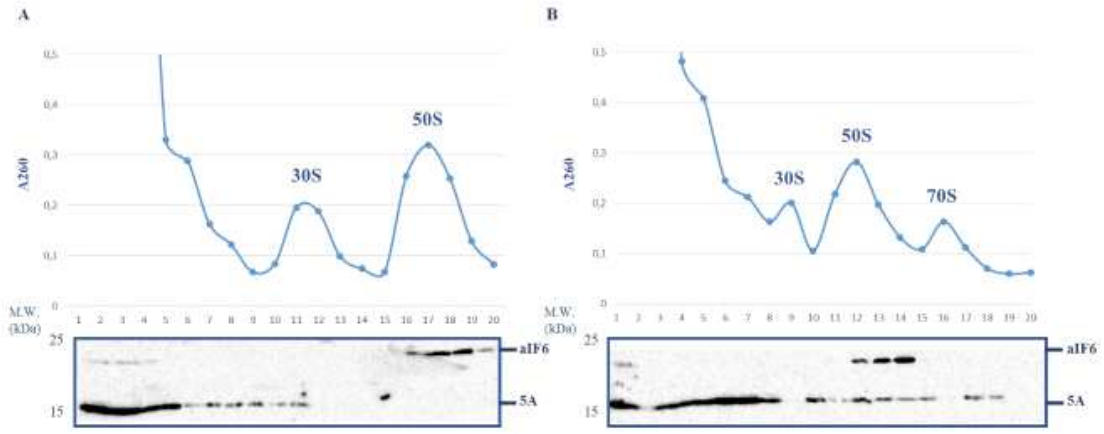


Fig.31 Fractionation of Sulfolobus solfataricus S30 extract and localization of Sso-aIF5A

(A) Density gradient fractionation of 500µg of total proteins Sso P2 S30 lysate fixed with formaldehyde as a control. (B) Density gradient fractionation of 500 µg of total proteins of Sso P2 S30 lysate programmed for translation and fixed with formaldeyde. Each fraction was separated on 15% SDS-gel and analyzed by immunoblot, using anti-aIF5A and anti-aIF6 antibodies.

Taken togheter these results are a first indication that an archaeal aIF5A protein participates in the process of protein synthesis, probably during the elongation phase of translation and the hypusine modification of aIF5A could mediate the interaction with ribosomes, as the eukaryotic eIF5A. Nevertheless, it would be interesting in future experiments to set up an *in vitro* translation assays in order to evaluate the activity of aIF5A in synthesis of proteins containing polyproline tracts, also using again both recombinants proteins (hypusinated and not) to evaluate the function of the hypusine residue more in detail.

3.7.2 A specific subset of RNA molecules are *in vivo* associated with Sso aIF5A

Several reports in the literature present evidence that IF5A is an RNA binding protein: a hypusine dependent RNA-binding activity of eIF5A was shown in yeast (Xu *et al*, 2001; Xu *et al,* 2004), while in an archaeal species it is proposed that aIF5A can not only bind, but also cleave RNA molecules. Wagner *et al* (2007) showed for the first time that aIF5A from *Halobacterium sp*. NRC-1 exhibits RNA cleavage activity, independent on post-translational modification but dependent on specific charged aa residues and occurred preferentially between adenine and cytosine in single stranded regions of RNA.
The halobacterial aIF5A was further shown to bind to RNAs *in vitro*, but in contrast to the RNase activity, the hypusine residue was required to stabilize RNA-protein complexes (Wagner *et al,* 2007).This is in agreement with the structure of both proteins as mentioned above, as aIF5A has a C-terminal domain with an Oligonucleotide binding fold, find in other proteins that is known to bind nucleid acids, as RNA chaperone, and this could be the portion of the protein that interact with nucleid acids, in particular RNA.
This observation prompted us to speculate that the archaeal IF5A fulfills different functions in the cell. First test was to investigate if *Sulfolobus solfataricus* aIF5A is associated *in vivo* to RNA substrates. In a deep sequencing analysis of RNA co-purified with recombinant aIF5A-C-His performed in a previous work (Bassani, PhD Thesis, 2017), a list of RNAs that copurifying with the protein was identified. Among these, there are mRNA and ncRNA.
Table 4 shows a group of mRNAs which have been selected for the highest aIF5A Reads and aIF5A TPM numbers, compared to the mock control:

Gene	Reads mock	Reads aIF5A	TPM mock	TPM aIF5A	Gene product
SSO-1773	39	944	0.6517	114.70	Multidrg-efflux transporter
SSO-3226	64	980	10	116.02	Fructose-bisphosphate aldolase
SSO-2184	3382	47164	381	3859.13	Cell division protein cdc6
SSO-0556	102	1397	23	230.37	CDP-diacylglycerol-glycerol-3-phosphate-3-phosphatidyltransferase
SSO-1896	178	2030	37	308.03	2-haloacid dehalogenase
SSO-0060	0	1293	0	129.38	UDP-N-acetylglucosamine-dolichyl-phosphate N-acetylglucosaminephosphotransferase (gnptA)
SSO-2413	91	1165	27	254.41	Hypotetical protein
SSO-2310	140	1791	23	212.04	Putative beta-lactamase
SSO-12199	162	2009	88	791.85	Putative MazG nucleotide pyrophosphohydrolase
SSO-0910	3904	47024	667	5845.50	Putative cell division protein

Table 4 mRNA co-purifying with Sso-aIF5A-C-His

Table 5 shows a group of ncRNA which have been selected for high 5A Reads and 5A TPM numbers, compared to the mock control:

Gene	Reads mock	Reads aIF5A	TPM mock	TPM aIF5A	Group
ncRNA26	3	214	0.2396	124.249	
ncRNA32	0	286	0	433.294	C/Dbox targets rRNA
SSOs04	0	156	0	270.106	snoRNA
ncRNA1	0	146	0	257.387	C/Dbox targets rRNA
ncRNA83	41	755	17.637	236.147	
ncRNA36	685	9505	327.415	3303.28	Complementary to tRNA-Pro
ncRNA2	130	1227	86.681	594.856	
ncRNA98	188	1496	278.566	1611.71	
ncRNA99	369	2400	332.489	1572.34	
ncRNA100	882	5312	712.848	3121.56	
ncRNA78	39	230	25.875	110.95	Trasposon related

Table 5 ncRNA co-purifying with Sso-aIF5A-C-His

In order to validate these data three of these RNAs were selected, including two mRNA and one ncRNA (that showed the highest number of reads in aIF5A-C-His eluate) to confirm their association with the native protein. The mRNAs chosen are SSO2184 and SSO0910.
SSO2184 gene is 1184 nt and encoding for a cell division control 6/orc1 protein homolog (cdc6-3) (KEGG database). SSO0910 gene is 780 nt and encoding for a putative cell division protein.
The gene coding for ncRNA98 is 90nt long but the function of this ncRNA is unknown.
To test if these RNAs are associated also with native protein *in vivo* and not only with recombinant protein, whose His-tag could induce a not specific interaction, a RNA immunoprecipitation (RIP) was performed. Native aIF5A was immunoprecipitated from *Sulfolobus solfataricus* post-ribosomal fraction (S100) using anti-aIF5A antibody and magnetic Dynabeads Protein G, that binds the antibodies via their Fc-region. Then immuno-complexes were isolated using a magnetic device, where the beads-Ab-aIF5A complex migrate to the side of the tube facing the magnet and allow for easy removal of the supernatant (unbound fractions). The complexs were washed, to avoid contamination of proteins bound to the tube wall, and eluted with an SDS-buffer to dissociate the aIF5A protein from the complex. As a control of specific aIF5A-binding a pre-immune serum was used, which is host rabbit serum extracted prior to immunization, supplied by the company that provided anti-aIF5A antibodies. Figure 32 shows the Western Blot analysis of different aIF5A-immunoprecipitation steps. Fig.32 (lane 7) highlights the presence of the native aIF5A only in the sample incubated with anti-aIF5A, and the absence of the protein in the control sample incubated with pre-immune serum (Fig.32, lane 8).

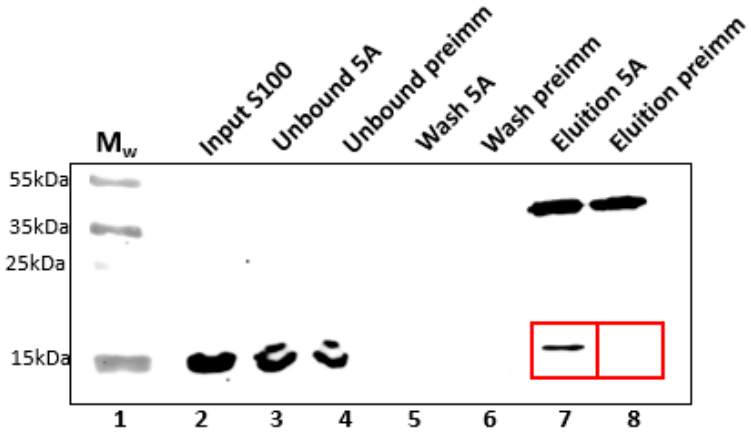


Fig.32 Native Sso-aIF5A immunoprecipitation from Sulfolobus solfataricus post-ribosomal fraction (S100).
IP with anti-aIF5A antibody and pre-immune serum and analysis of immunocomplexes by western blotting, using anti-aIF5A. Lane 1: molecular weight protein marker, lane 2: input of native aIF5A in S100, lane 3-4: unbound fractions, lane 5-6: last wash fractions, lane 7-8: elution fractions. Native aIF5A in the IP anti-aIF5A elution sample and the absence in the IP pre-immune serum elution sample are shown with red boxes.

RNAs directly bound to native aIF5A were than extracted from both this elution fractions, to ensure that the extracted RNAs were directly related to aIF5A and not to non-specific interactions with the magnetic beads, and DNase treated. The RNAs were retrotrascribed and cDNAs were used as a template for PCR amplification using specific primers that amplify internal sequences of mRNA 2184, mRNA 0910, nc98 and 2508 mRNA (another Sso mRNA not present in a list of RNA co-purified with aIF5A-C-His). The reactions were carried out with and without Reverse Transcriptase, to exclude any DNA contamination. Figure 33 shows that there is an amplification band only in PCR-reactions with oligonucleotides that amplified mRNA2184, mRNA0910 and ncRNA98, choosen from the list of RNAs co-eluted with aIF5A (Fig.33 lane 3), and this confirms that this three RNAs are associated to the aIF5A protein also *in vivo*.

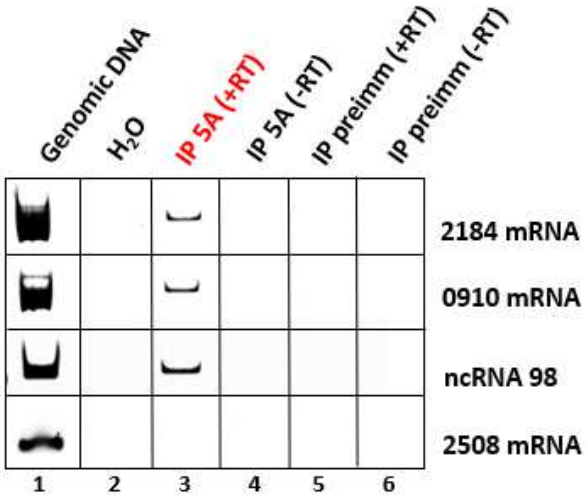


Fig.33 Reverse-Transcriptase-PCR with primers that amplified 2184, 0910, 2508 mRNA and ncRNA98. Lane 1: Sso P2 genomic DNA amplification (control for PCR), lane 2: sample with H_20, to exclude any contamination of genomic DNA in water used for PCR, lane 3-4: RNA- coimmunoprecipitated with native aIF5A (IP 5A), lane 5-6: RNAs in control samples of immunoprecipitation with preimmune serum. (IP preimm). The reactions were carried out with (+RT) and without (-RT) Reverse Transcriptase, to exclude any DNA contamination.

3.7.3 Sso aIF5A exhibits *in vitro* mRNA ribonucleolytic activity, which does not require hypusination

Based on the studies described above and on our findings of an interaction between aIF5A and RNA molecules, next we questioned whether aIF5A is also endowed with RNase activity as discovered in *Halobacterium sp.* The RNA substrates aIF5A-bound identified, were *in vitro* transcribed and used in a *in vitro* RNA cleavage assay to test the RNA degradation activity of both recombinant unmodified proteins purified from *E. coli* (N-His-aIF5A) and the recombinant hypusinated aIF5A purified from *Sulfolobus solfataricus* (aIF5A-C-His). The two proteins were incubated at 65°C for 30 min in presence of these different length RNA substrate: mRNA 2184 (1184 nt), mRNA 0910 (780 nt), ncRNA 98 (90 nt), in a ratio RNA/protein 1:4. Every 5 minutes an aliquot was withdrawn and analyzed by denaturing Urea-PAGE. Experiments were repeated at least three times and representative results are shown in Figure 34.

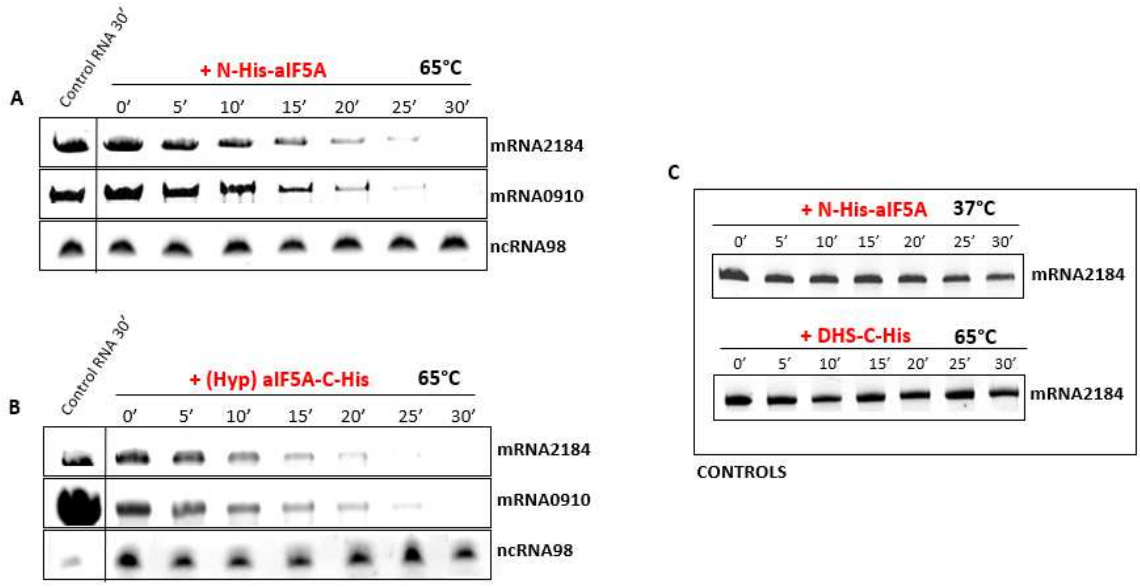


Fig.34 Sso aIF5A (N-His-aIF5A from E. coli and aIF5A-C-His from Sso) show ribonucleolityc activity.
(A) Degradation of mRNA2184, mRNA0910 and ncRNA98 by unmodified N-His-aIF5A from *E.coli* is shown on a native polyacrylamide gel 8% in a RNA/protein ratio 1:4 at 65°C. (B) Degradation of mRNA2184, mRNA0910 and ncRNA98 by hypusinated aIF5A-C-His from *Sso* is shown on a native polyacrylamide gel 8% in a RNA/protein ratio 1:4 at 65°C. (C) Control assays with mRNA2184 and N-His-aIF5A from *E.coli* at 37°C and mRNA2184 and Sso-DHS-C-His at 65°C, both on a native polyacrylamide gel 8% in a RNA/protein ratio 1:4

These results represent a clear indication that both unmodified (Fig.34A) and hypusinated (Fig.34B) aIF5A has a ribonucleolytic activity towards the mRNA substrates (2184, 0910), in a hypusine independent way. Curiously, this it is not the case for the small ncRNA98, which is not degraded by aIF5A, so probably the cleavage activity depends on the RNA substrate.

For each reaction a control sample of RNA without protein was incubated for 30 min at 65°C, to prove the stability of these RNAs at high temperature. Further evidence that RNA degradation is due to the aIF5A activity, other controls were fulfilled (Fig.34C). The assay was carried out with mRNA 2184 with N-His-aIF5A at 37°C, to exclude the presence of *E. coli* RNases in protein preparation, and mRNA 2184 with another recombinant protein from *Sulfolobus solfataricus* expressed and purified in *E. coli* (the DHS protein): in both controls no degradation of the RNA substrate was observed.

Among the aIF5A co-purifying RNAs neither a consensus cleavage sequence could be identified, nor a conserved ribonucleolytic domain in Sso aIF5A protein sequence. However, it was shown that specific amino acid exchanges in the protein sequence, affect the

ribonucleolytic activity of aIF5A in *Halobacterium sp*. NRC-1 (Wagner *et al*, 2007). In particular the ribonucleolytic activity of aIF5A was reduced when amino acids at position 9 (N-terminal domain), amino acids at position 72/73 (hinge region), or amino acids at position 117 or 122/123 (C-terminal domain), were exchanged. From the alignment of aIF5A protein sequences from *Halobacterium sp.* NRC-1 and from *Sulfolobus solfataricus* P2 the three residues in the C-terminal domain are conserved (Figure 35, residues highlighted by red asterisks) and they might be involved in the RNA cleavage. Therefore, this could explain why the hypusine residue is unnecessary for ribonucleolytic activity of aIF5A.

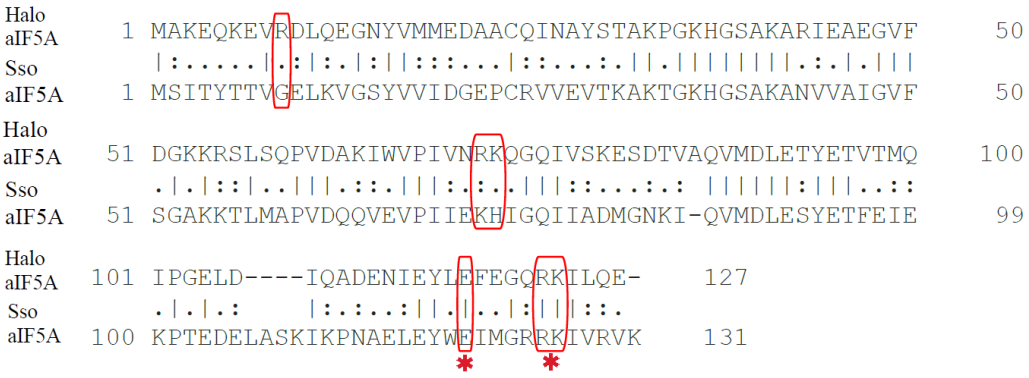


Fig.35 Conservation of aIF5A residues involved in RNA cleavage. Protein sequences alignment of aIF5A from *Halobacterium sp* NRC-1 (HaloaIF5A) and from *Sulfolobus solfataricus* P2 aIF5A (Sso aIF5A).

3.7.4 Sso aIF5A shows endonucleolytic activity

The data highlight the possible role of the free form (not ribosome associated) of aIF5A in RNase cleavage *in vitro*. Next, we wondered if this archaeal protein displays endoribonucleolytic or esoribonucleolytic activity. To obtain the RNA degradation pattern, a cleavage assay was performed, with a radiolabelled 2508sh mRNA substrate, an RNA model used for other similar works in *Sulfolobus* (Hasenöhrl *et al*, 2011). The RNA was incubated for 24' at 65°C in presence of N-His-aIF5A (from *E. coli*) in a molar ratio of 1:4. Every two minutes an aliquote was withdrawn and loaded on 8M Urea-Polyacrylamide 20% gel. As shown in the figure 36A, the full length of mRNA (253 nt) decrease over time, consistent with the degradation activity of aIF5A, while a band of around 80 nt long seems to accumulate. If we look at the predicted structure of this substrate (Fig.36B), at this position there is a single strand region and bulges that could be the targets of aIF5A. As a matter of

fact, the ribonucleolytic activity of *Halobacterium* aIF5A seems to be restricted to single-stranded CA bonds (Wagner *et al*, 2007). Although the mechanism of cleavage of Sso aIF5A remains to be tested, from the degradation pathway observed as well as from the prediction of the structure of RNA substrate used in the *in vitro* assay, we can assume that the cleavage preferentially occurs within single stranded regions or bulges and might involves GU or GA bonds.

So, presents data suggest that the protein acts as an endonuclease, and does not require hypusine for his activity. The mechanism of cleavage of aIF5A however remains to be tested with other RNA substrates to understand if this activity is specific or not.

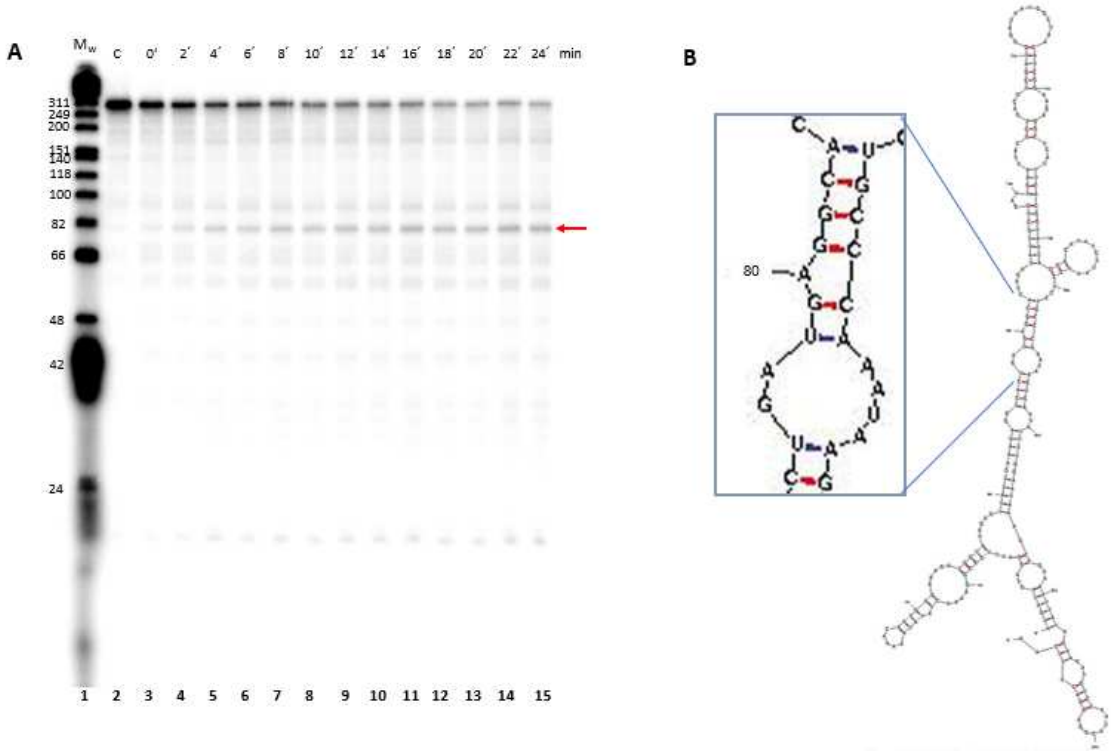


Fig. 36 S. solfataricus aIF5A is an endoribonuclease. (A) Lane 1, φX174 DNA/HinfI 5´-end radioactively labeled marker. Lane 2, incubation of 2508sh RNA for 22 min at 65°C, in the absence of aIF5A. Lanes 3-14, Incubation of 2508sh RNA, 5´-end radioactively labeled, with N-His-aIF5A for 0 to 22 min at 65°C. The arrow indicates the accumulation of a breakdown product. (B) Structure prediction of Sso 2508sh mRNA obtained from mfold web server.

3.7.5 Hypusine-dependent ncRNA binding of Sso aIF5A

It was demonstrated that aIF5A is able to cleave at 65°C the two mRNAs substrate tested in the *in vitro* degradation assays, independently of hypusination, but curiosly the ncRNA98 remained intact after incubation with both recombinant aIF5A proteins. Anyhow it was evinced that this ncRNA98 was associated to the native protein *in vivo,* hence we wondered if aIF5A could displays also RNA-binding capacity *in vitro*, as the eukaryotic eIF5A and *Halobacterium sp*. aIF5A.

To investigate if aIF5A could form RNA-protein complex with this small ncRNA, an electrophoretic mobility shift assay (EMSA) was employed, using again the unmodified and the hypusinated form of aIF5A.

5'-radiolabelled-nc98 was incubated with increasing concentration of aIF5A proteins at 65°C for 20 min and loaded on 8% polyacrylamide gel.

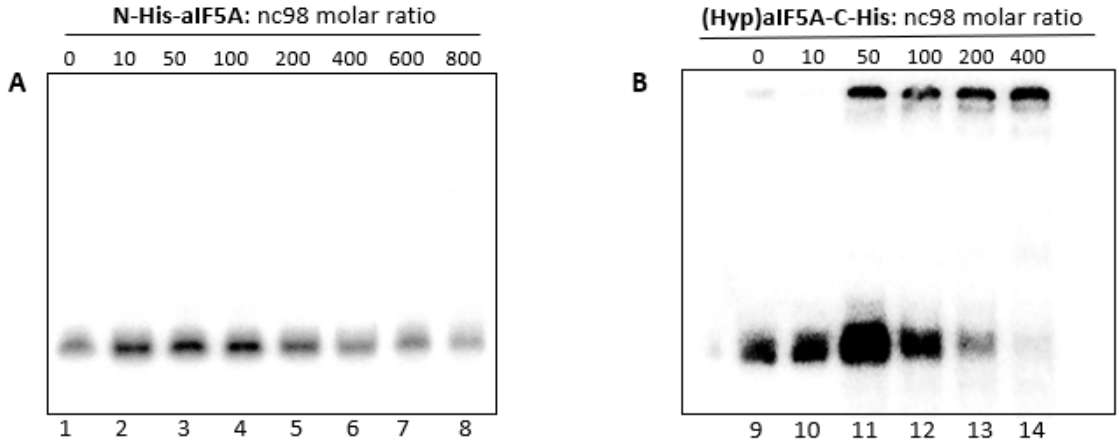


Fig.***37 Binding of the ncRNA98 to hypusinated aIF5A-C-His***. EMSA assays were performed with the 5'-labelled ncRNA98 substrate and the unmodified N-His-aIF5A (from *E. coli*) (A) and with hypusinated aIF5A-C-His (form Sso) (B) in a molar ratio protein:RNA of 0:0, 10:0, 50:0, 100:0, 200:0, 400:0, 600:0, 800:0.

As shown in figure 37, starting from a molar ratio of ncRNA98/aIF5A of 0:50 a partial mobility shift of the RNA was obtained with hypusinated aIF5A-C-His (Fig.37B, lane 11), whereas no significant shift was obtained with the unmodified protein N-His-aIF5A, even when the protein was added in 800-fold molar excess (Fig.37A, lane 8).

Hovewer, it should be notice that the two proteins used in this assay differ, as well as the presence of hypusine, also for the different position of His-tag (N- and C-terminal), which

could influence the binding with the ncRNA. In order to validate this result, this experiment should be repeated using the native protein from Sso, which is hypusinated and is devoid of His tag.

Anyhow, this preliminary result shows that the protein could not only cleave, but also bind some RNA molecules, consistent with findings reported in yeast and *Halobacterium sp*. by Xu and Wagner; in this case Sso aIF5A is able to interact with a Sso non-coding RNA 98, of which unfortunately nothing is known.

3.8 Sso aIF5A is expressed throughout the growth curve

As a final but still preliminary investigation, considering the multifunctionality of the protein, we want to evaluate if the expression of the protein changes during growth phases and whether it is affected by temperature alterations. aIF5A levels at different growth phases were monitored by Western Blot using Sso anti-aIF5A antibodies. This antibody detected aIF5A at around 15 kDa, exactly the expected theoretical mass predicted on Expasy tool for native protein (14491.86 Da). To get some insight into the regulation of aIF5A, we asked whether the protein was differentially expressed during the cell-cycle or in response to enviromental changes. First, we evaluated the amount of aIF5A present in different phases of cell growth: early exponential (0.2 OD_{600}), mid exponential (0.4 OD_{600}), late exponential (0.8-1.3 OD_{600}), and stationary phase (1.8 OD_{600}). As shown in Figure 38B similar levels of aIF5A protein were detected for up to 34 hours of growth (1.82 OD_{600}), until the stationary phase, suggesting that the expression of the factor is not cell-cycle regulated. Next, *S. solfataricus* cells growth to late exponential phase (0.8 OD_{600}) at the normal temperature of 75°C were subjected to both cold-shock and heat-shock, by transferring the cultures for 30 min at 60°C and 90°C, respectively. But, also in this case, no substantial differences in aIF5A abundance were detected. These preliminary data suggest that the protein is very stable and always expressed, also under unfavorable stress conditions (Fig 38C).

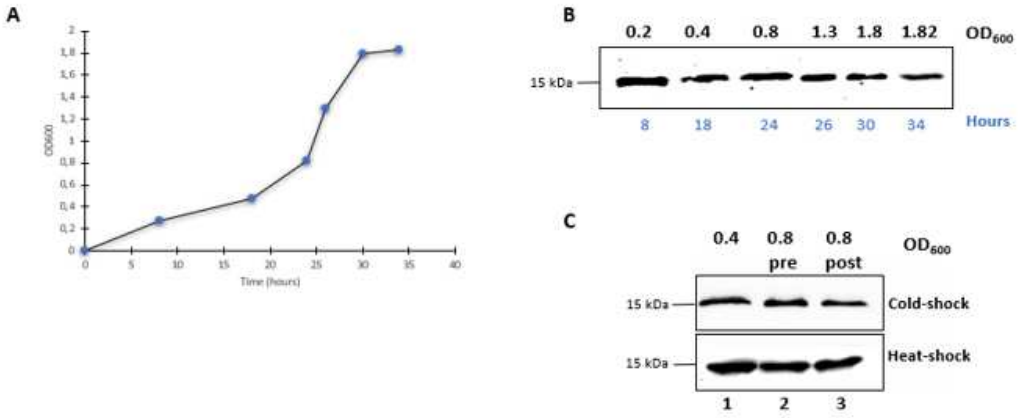


Fig.38 Expression of Sso aIF5A in the cell cycle and upon thermal stress. (A) *S. solfataricus* P2 was grown at 75°C. 15ml of culture at 0.2 OD (8h), 0.4 OD (18h), 0.8 OD (24h), 1.3 OD (26h), 1.8 OD (30h) and 1.82 (34h), were harvested by centrifugation at 4000 rpm for 20min at 4°C. (B) 20 µg of total protein from each lysate at different densities, were elecrophoresed and probed with anti-aIF5A antibody. (C) Effect of thermal change on aIF5A levels. Western Blot analysis with anti-aIF5A of 20 µg of total protein from mid exponential phase culture (0.4 OD) lane1, late exponential phase culture (0.8 OD) lane 2, and late exponential phase colture (0.8 OD) after 30 min at 60°C (high panel) or 90°C (low panel).

4. Discussion

The work presented in this thesis provides new insights on the archaeal translation factor aIF5A, in the Crenarchaeota *Sulfolobus solfataricus*.

Translation is a very important step in the expression of genetic information in all three domains of life. It is an extremely conserved process and an example of this high degree of conservation is represented by translation factors. Some are homologous in all three domains of life, among these a/eIF5A in Archaea and Eukarya and EF-P in Bacteria (Benelli *et al*, 2009a).

The translation machinery has been extensively studied in Bacteria and Eukaryotes, whereas the archaeal one is still poorly characterized. Archaeal microorganisms, omnipresent in Earth's ecosystems, constitute a unique domain of life with features at the frontier of Bacteria and Eukarya. Hence, the elucidation of fundamental cellular processes as translation could, on one hand, increase knowledges about this kingdom, currently still very limited and, on the other hand, contribute, from an evolutionary point of view, to understand the origin of life on earth and how more complex organisms, as eukaryotes, have evolved. Current evolutionary model suggest indeed that Eukaryotes originated from Archaea instead of being a sister lineage (Lyu *et al*, 2017). In addition, the close phylogenetic relationship between Archaea and Eukarya point out the prospective of using archaeal organisms also as useful model system to study eukaryotic complex information processing systems, in a simpler background easier to interpret. Furthermore, being the archaeal group adapted to live in many extreme conditions, the characterization of their proteins, in particular from hyperthermophiles as *Sulfolobus*, could be useful in biotechnological applications (Cavicchioli, 2011). All these observations greatly stimulated our interest in these organisms and we focused our research on the translation apparatus of the model organism *S. solfataricus.*

In particular the aim of this work was to obtain a structural and functional characterization of one of the universally conserved translation factors the *S. solfataricus* aIF5A. Its function and properties have been studied to some extent in Eukarya and in Bacteria but remain somewhat elusive. The homologous proteins, eIF5A in Eukarya and EF-P in Bacteria, are known to act as translation factors, in particular promoting the synthesis of proteins containing specific amino acid sequences such as consecutive polyproline stretches (Gutierrez *et al*, 2013). These motifs in fact cause the stall of ribosomes because of the constrained geometry of the cyclic side chain of proline, but eIF5A/EF-P localize between

the P and E site of the ribosome and interact with the end-terminus of tRNA, rRNA and ribosomal proteins (Melnikov *et al*, 2016), rescuing the polyPro ribosomes blocked. Essential for this function is the unique post-translational modification typical of these factors: the hypusine for eIF5A and lysyl-lysine for EF-P. Very little is known about aIF5A in the archaeal kingdom. Concerning its post-translational modification, the distribution of hypusine residue in Archaea is heterogenous. Several anaerobic species contain deoxyhypusine, other selected species like *Sulfolobus acidocaldarius, Halobacterium cutirubrum* and *Thermoplasma acidophilum* contain hypusine (Schumann *et al*, 1990). Curiously only the gene for the first enzyme involved in the modification pathway (deoxyhypusine synthase or DHS) has been identified, whereas the second enzyme (deoxyhypusine hydroxylase or DOHH) has never been found in any archaeal genomes to date.

Since nothing is known concerning the role of archaeal aIF5A and the hypusination pathway in these organisms, in order to fill the gap a series of experiments were performed to shed light on this interesting archaeal protein in hyperthermophilic *Sulfolobus solfataricus*.

In the first part of this work we show for the first time that the native protein, isolated from Sso lysate, undergoes the hypusine modification as the eukaryotic protein. The presence of this post-translational modification was proved first by Western Blotting analysis, with an antibody recognizing the hypusine-containing peptide, and then by mass spectrometry (figure 17). The mass analysis confirmed the identity of hypusine and highlighed that this residue is attached to lysine 36 of Sso aIF5A.

The presence of hypusine modification at the N-terminal position of aIF5A immediately suggests a strong similarity between Archaea and Eukarya. Nevertheless, the question about how this hypusine residue is synthesized remains unanswered because of the lack of the second enzyme.

We, therefore, focused our attention to the structural characterization of both aIF5A proteins and the first enzyme of modification, DHS. The overproduction of the recombinant aIF5A protein was performed both in *E. coli* and directly in *S. solfataricus*, in order to obtain both forms of the protein (the unmodified N-His-aIF5A and the hypusinated aIF5A-C-His respectively). Mass spectrometry analysis of the aIF5A-C-His expressed in Sso confirmed that the recombinant protein becomes hypusinated, as the native one (figure 18).

We started from the quaternary structure of aIF5A: previous data had proposed that the homologous eukaryotic protein assumes a dimeric conformation in solution and the molecular envelope of the dimer is similar to the bacterial EF-P monomer (Dias *et al*, 2013). In addition, also the crystal structure of *Methanococcus jannaschii* aIF5A suggests the formation of a dimer (Kim *et al*, 1998). We performed SAXS (small-angle X-ray scattering) experiments at the Diamond synchrotron light source (UK), analyzing different concentrations of the unmodified recombinant protein in solution at high temperature (63°C), to mimic as much as possible the condition of its physiological state. From the analysis of SAXS curves (figure 20) we deduced the monomeric status of N-His-aIF5A in solution, and using the Quafit software, it was possible to optimize 10 conformation of the protein, in which the N-terminal and C-terminal domains of this conformers are more rigid, and they differ for the flexibility connected by more flexible parts (figure 22), in agreement with the crystal structure of the homologous proteins in Eukarya and in other Archaea.

Gel-filtration analysis confirmed the monomeric state of the protein and showed that this state is independent of the presence of hypusine (figure 23), in contrast with what proposed in eukaryotes (Gentz e*t al*, 2009).

We proceeded with the structural characterization of the deoxyhypusine synthase (DHS) in Sso. While cloning the DHS enzyme in *E. coli*, we observed that the recombinant enzyme was soluble only when the His-tag was located at the C-terminal position (figure 19). The eukaryal DHS is a tetramer and consists of two dimers associated by their N-terminal domains, with the two active sites at the interface between them (Liao *et al*, 1998), therefore our results could be a first indication that also Sso DHS forms a tetramer with a structural organization similar to the eukaryotic protein. In fact, the insolubility of the N-terminal-tagged Sso DHS protein could be explained by the fact that the his-tag interferes with the proper folding of the enzyme. The structural analysis with SAXS of the recombinant DHS-C-His revealed that this speculation is correct, since we showed that also the archaeal DHS is a tetramer, as the eukaryotic counterpart (figure 24).

The presence of hypusine and the tetrameric conformation of DHS enzyme in Sso, exactly as in Eukarya, suggest that DHS enzyme is implicated in the modification of aIF5A. To verify this, we first investigate if these two recombinant proteins, aIF5A and DHS, interacted in vitro to form a complex. Employing SAXS analysis (figure 25), gel-filtration (figure 26) and non-denaturing gel electrophoresis (figure 27), we demonstrated that the DHS-C-His

enzyme and the N-His-aIF5A precursor form a stable complex *in vitro* at 65°C. Moreover, the interaction between the two proteins is independent on the presence of the other components of the enzymatic reactions, NAD^+ and spermidine, as demonstrated for human DHS (Lee *et al,* 1999). Therefore, these results hinted that Sso DHS is involved in the modification of aIF5A protein but the hypusination reaction in Archaea remains obscure. Interestingly, in *Trichomonas vaginalis* it was proposed that the enzyme Tv-DHS, in addition to its deoxyhypusine synthase activity, is endowed with a hydroxylase activity for the formation of hypusine in the absence of DOHH (Quintas-Granados *et al*, 2015). This observation suggests to test if this feature could be shared also by Sso DHS. Our results are in disagreement with this study, as the *in vitro* hypusination assay unveiled that Sso DHS is able to perform the sole deoxyhypusination of the protein, although its activity is not efficient (figure 28). These data are contrariwise consistent with analysis of eukaryotes DHS activity, that performs only the first step of hypusine synthesis in eIF5A precursor, in a reversible reaction (Park *et al*, 2003). This could explain the low amount of deoxyhypusine catalyzed by Sso DHS, compared to the total unmodified protein (figure 28), because in absence of the second enzyme of modification, unknown to date, the reaction reverses quickly in the opposite direction.

Future efforts will be employed to disclose the second enzyme of hypusine modification in order to clarify the hypusination pathway in *S. solfataricus*.

In our previous work we identified the Sso aIF5A interactome, but among these no proteins with sequence similarity or functional analogy with DOHH were identified (Bassani *et al*, 2018).

The second part of this work was dedicated to a functional characterization of the archaeal translation factor aIF5A. We propose that *S. solfataricus* aIF5A could be a multitasking protein: in fact, on one hand we provide the first evidence that the protein plays a role in the translation process but, on the other hand, it might be endowed with different functions related to RNA-processing pathways.

Based on our studies, the native protein appears to be abundant either in the post-ribosomal fractions (S100 and High Salt Wash) and also associated with 70S ribosomes (figure 29). aIF5A seems to interact *in vitro* with the Sso 50S ribosomes, probably through the hypusine residue (figure 30). In addition, the behavior of the protein during protein synthesis was investigated immunolocalizing the native protein on a Sso lysate “programmed for

translation", by addition of exogenous mRNA, fixed with formaldehyde and fractionated on sucrose density gradient. The protein is mainly present as a free form in a translationally silent Sso extract, but it becomes ribosome bound, being detected in the 30S, 50S and 70S fractions, when a *Sulfolobus* cell lysate has been programmed for translation (figure 31). This represents the first indication of the involvement of an archaeal aIF5A in the translation process, confirming an evolutionary conserved role of this translation factor.

As mentioned before, the function of eukaryal eIF5A and bacterial EF-P is strictly related to a translation of proteins with consecutive proline motifs (Schuller *et al*, 2017). However, a genome-wide analysis has shown that these proteins with polyproline stretches are not so common in Archaea (Mandal *et al*, 2014), so probably to limit the future researches focusing only on translation of these polyproline proteins would be inaccurate.

In this work we proposed that aIF5A protein is a multifunctional protein.

A deeper investigation of the possible role of the free form (not ribosome associated) of aIF5A revealed that the recombinant hypusinated aIF5A-C-His is associated with some RNA substrates (mRNA and ncRNA), that co-purified with the protein (Bassani, PhD thesis, 2017). We validated this data, showing also for the native aIF5A *in vivo* an interaction with three of those Sso RNA substrates, that include two mRNA (2184 and 0910) and one ncRNA (nc98) (figure 33). Starting from this data, we found that the aIF5A protein (modified and not) displays an RNase activity *in vitro* for this mRNA molecules at 65°C (figure 34). Our results are in agreement with previous studies that show an RNA cleavage activity of aIF5A from *Halobacterium sp.* NRC-1 (Wagner *et al,* 2007). Although a conserved ribonucleolytic domain in aIF5A has never been identified, Wagner and coworkers show that changing specific amino acids reduced the ribonucleolytic activity of Halo-aIF5A. When we compared amino acids sequences from *Halobacterium sp.* NRC-1 and *S. solfataricus* P2 aIF5A, three of these residues localized in C-terminal domain are conserved (figure 35), and might be involved in the RNA cleavage. Furthermore, this could explain why the hypusine residue is not required for this cutting activity, as this modification places in N-terminal position of the protein.

Additionally, this study shows that Sso aIF5A is an endonuclease.

Using a 5'-end-radiolabelled natural mRNA from *S. solfataricus* we obtain a pathway of degradation that suggests, together with the prediction of the structure of this RNA (figure 36), that the cleavage occurs within single stranded regions or bulges and might involves

GU or GA bonds. This behaviour has also been reported for *Halobacterium sp.* aIF5A, whose ribonucleolytic activity is restricted to single-stranded regions, in particular to CA bonds (Wagner *et al*, 2007).

Further studies are necessary to clarify the mechanism of cleavage of Sso aIF5A, testing its activity with other RNA substrates, and to investigate the role of the protein in RNA-processing pathways in *Sulfolobus solfataricus*.

Besides, information on mechanisms by which RNA species reach their mature forms and associated RNA-modifying proteins are still fragmentary in Archaea. Several endonucleases have been identified from the information gathered from Bacteria and Eukarya, as the orthologue of RNase P, the splicing endonuclease End A, aRNase Z, aPelota, members of the ß-CASP ribonuclease family (aCPSF1, aCPSF2), aNob1 and CRISP Cas6 (Clouet-d'Orval *et al*, 2018). In this framework, aIF5A could be a new RNase candidate that, in complex with other components, partecipates in RNA processing and decay paths in Sso.

Finally, we propose that aIF5A displays a binding activity towards the other RNA tested in this work as well (ncRNA98). This ncRNA interacts with the protein *in vivo* (figure 33), but aIF5A is unable to cleave this substrate (figure 34).

The finding that IF5A is a binding protein was proposed either in Eukarya, as in previous studies was demonstrated that eIF5A can bind some RNA molecules (Xu *et al*, 2001), and also in *Halobacterium sp.* (Wagner *et al*, 2007). In both cases this activity is mediated by the hypusine residue. This supports our hypothesis that aIF5A from Sso binds a non-coding RNA substrate, in hypusine-dependent manner, as the 5'-end-labelled ncRNA98 is restrained in an electromobility shift assay only when the hypusinated-aIF5A is employed in the assay (figure 37).

In eukaryotes, the RNAs bound to eIF5A shares conserved motifs (UAACCA and AAAUGU) (Xu *et al*, 2004). However, a consensus cleavage sequence has not been identified in the list of RNAs that co-purified with the Sso protein. The binding ability may be due to the structural conformation of the protein: crystal structures of archaeal homologous of translation factor 5A reveals that the C-terminal domain has an oligonucleotide binding fold (OB-fold), found in other proteins that is known to bind nucleic acids, as CspA, the major cold shock protein of *E. coli*, which is known to be an RNA chaperone (Jiang *et al*, 1997). Thus, it is possible that the *in vitro* RNA binding of Sso aIF5A is mediated by this fold and the charged hypusine residue is important to stabilize the RNA-

protein complex formation. This could be the reason why aIF5A-RNA interaction only be shown for the hypusinated protein *in vitro,* because the binding of the unmodified protein would only be transient and not stable enough to be seen in gel shift experiments.

Further works are needed to understand if this binding activity of aIF5A targets structural characteristic elements of some RNAs or if the interaction with this non-coding RNA has a functional meaning. Although the function of ncRNA98 is not known, the non-coding RNAs play a lot of roles in the cells and in *S. solfataricus* participates in many different biological processes including transcriptional regulation, mRNA processing and tRNA and rRNA regulation (Tang *et al* 2005).

Finally, considering the multiple functions proposed for the aIF5A protein and the structural similarity in the C-terminal domain with the cold shock proteins, we therefore analyzed in a preliminary experiment the expression of the protein during the growth phases and if it is affected by temperature changes, as cold and heat shock. Our results show that the archaeal aIF5A is always expressed during all the growth phases, also under unfavorable temperature changes (figure 38).

Concluding, the *Sulfolobus solfataricus* aIF5A probably owns a functional dualism on RNA-processing activities: the protein shows a cleavage activity for some mRNA substrates and on the other hand displays an RNA binding activity. These two opposing activities could be regulated by the post-translational modification status of the protein (hypusinated vs unhypusinated) and by its interaction with different protein partners.

Therefore, the aIF5A protein could play a role in a wider scenario, in which the protein is part of a more complex system, involved in regulation and metabolism of different type of RNAs.

5. References

Aoki H, Dekany K, Adams SL, Ganoza MC. (1997) The gene encoding the elongation factor P protein is essential for viability and is required for protein synthesis. J Biol Chem. 272(51):32254-9.

Aoki H, Xu J, Emili A, Chosay JG, Golshani A, Ganoza MC. (2008) Interactions of elongation factor EF-P with the *Escherichia coli* ribosome. FEBS J. 275: 671–81.

Balibar CJ, Iwanowicz D, Dean CR. (2013) Elongation factor P is dispensable in *Escherichia coli* and *Pseudomonas aeruginosa*. Curr Microbiol. 67(3):293-9

Bartig D, Lemkemeier K, Frank J, Lottspeich F, and Klink F (1992) The archaebacterial hypusine-containing protein: structural features suggest common ancestry with eukaryotic translation initiation factor 5A. Eur. J. Biochem. 204: 751–758.

Bartig D, Schümann H, Klink F. (1990) The unique posttranslational modification leading to deoxyhypusine or hypusine is a general feature of the archaebacterial kingdom. System. Appl. Microbiol. 13: 112–116.

Bassani F. (2017) PhD thesis. New functions for an archaeal translation factor. Università Politecnica delle Marche.

Bassani F, Romagnoli A, Cacciamani T, Amici A, Benelli D, Londei P, Märtens B, Bläsi U, La Teana A. (2018) Modification of translation factor aIF5A from *Sulfolobus solfataricus*. Extremophiles 22: 769-780.

Behshad E, Ruzicka FJ, Mansoorabadi SO, Chen D, Reed GH, Frey PA. (2006) Enantiomeric free radicals and enzymatic control of stereochemistry in a radical mechanism: the case of lysine 2,3-aminomutases. Biochemistry 45(42):12639-46

Benelli D, Londei P (2007) In Vitro Studies of Archaeal Translational Initiation. Methods in Enzymology 430: 79-109.

Benelli D, Londei P. (2009) Begin at the beginning: evolution of translational initiation. Res Microbiol. 160: 493–501.

Benelli D, Maone E, Londei P. (2003) Two different mechanisms for ribosome/mRNA interaction in archaeal translation initiation. Mol. Microbiol. 50: 635-643.

Benelli D, Marzi S, Mancone C, Alonzi T, la Teana A, Londei P. (2009) Function and ribosomal localization of aIF6, a translational regulator shared by archaea and eukarya. Nucleic Acids Res. 37: 256e267.

Benne R, Brown-Luedi ML, Hershey JWB. (1978) Purification and characterization of protein synthesis initiation factors eIF-1, eIF-4C, eIF-4D, and eIF-5 from rabbit reticulocytes. J Biol Chem. 253: 3070–3077.

Blaha G, Stanley RE, Steitz TA. (2009) Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science 325(5943):966-70

Brock TD, Brock KM, Belly RT, Weiss RL. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch. Mikrobiol. 84: 54–68.

Bullwinkle TJ, Zou SB, Rajkovic A, Hersch SJ, Elgamal S, Robinson N, Smil D, Bolshan Y, Navarre WW, Ibba M. (2013) (R)-β-Lysine-modified elongation factor P functions in translation elongation. J Biol Chem. 288: 4416–23.

Caraglia M, Park MH, Wolff EC, Marra M, Abbruzzese A. (2013) eIF5A isoforms and cancer: two brothers for two functions? Amino Acids. 44(1):103-9

Cavicchioli R. (2011) Archaea, timeline of the third domain. Nature Reviews 9:51-61

Chatterjee I, Gross SR, Kinzy TG, Chen KY. (2006) Rapid depletion of mutant eukaryotic initiation factor 5A at restrictive temperature reveals connections to actin cytoskeleton and cell cycle progression**.** Mol Genet Genomics. 275(3):264-76

Choi S, Choe J. (2011) Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 Å resolution. Proteins 79(5):1688-93.

Clement PM, Henderson CA, Jenkins ZA, Smit-McBride Z, Wolff EC, Hershey JW, Park MH, Johansson HE. (2003) Identification and characterization of eukaryotic initiation factor 5A-2. Eur J Biochem. 270(21):4254-63

Clouet-d'Orval B, Batista M, Bouvier M, Quentin Y, Fichant G, Marchfelder A, Maier LK. (2018) Insights into RNA-processing pathways and associated RNA-degrading enzymes in Archaea. FEMS Microbiology Reviews, 42: 279-613.

Condo I, Ciammaruconi A, Benelli D, Ruggero D, Londei P. (1999) Cis-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. Mol. Microbiol. 34: 377–384.

Coureux PD, Lazennec-Schurdevin C, Monestier A, Larquet E, Cladièr L, Klaholz BP, Schmitt E, Mechulam Y. (2016) Cryo-EM study of start codon selection during archaeal translation initiation. Nature Communications 7: 13366.

Cox J, Mann M. (2008) MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26:1367–1372

Dever TE, Gutierrez E, Shin BS. (2014) The hypusine-containing translation factor eIF5A. Crit Rev Biochem Mol Biol. 2014 Sep-Oct;49(5):413-25

Dias CA, Garcia W, Zanelli CF, Valentini SR. (2013) eIF5A dimerizes not only in vitro but also in vivo and its molecular envelope is similar to the EF-P monomer. Amino Acids. 44(2):631-44

Gäbel K, Schmitt J, Schulz S, Näther DJ, Soppa J. (2013) A comprehensive analysis of the importance of translation initiation factors for *Haloferax volcanii* applying deletion and conditional depletion mutants. PLoS One. 8(11): e77188.

Ganoza MC, Aoki H. (2000) Peptide bond synthesis: function of the efp gene product. Biol Chem. 381(7):553-9

Gentz PM, Blatch GL, Dorrington RA. (2009) Dimerization of the yeast eukaryotic translation initiation factor 5A requires hypusine and is RNA dependent. The FEBS Journal, 276(3), 695–706.

Glick BR, Chladek S, Ganoza MC. (1979) Peptide bond formation stimulated by protein synthesis factor EF-P depends on the aminoacyl moiety of the acceptor. Eur J Biochem. 97: 23–28.

Glick BR, Ganoza MC (1975) Identification of a soluble protein that stimulates peptide bond synthesis. Proc Natl Acad Sci U S A. 72(11): 4257-60.

Gosslau A, Jao DL, Butler R, Liu AY, Chen KY. (2009) Thermal killing of human colon cancer cells is associated with the loss of eukaryotic initiation factor 5A. J Cell Physiol. 219(2):485-93

Grill S, Gualerzi CO, Londei P, Bläsi U. (2000) Selective stimulation of translation of leaderless mRNA by initiation factor 2: evolutionary implications for translation. EMBO J. 19:4101-4110.

Gutierrez E, Shin BS, Woolstenhulme CJ, Kim JR, Saini P, Buskirk AR, Dever TE. (2013) eIF5A promotes translation of polyproline motifs. Mol Cell. 51: 35–45.

Hanauske-Abel HM, Park MH, Hanauske AR, Popowicz AM, Lalande M, Folk JE. (1994) Inhibition of the G1-S transition of the cell cycle by inhibitors of deoxyhypusine hydroxylation. Biochim Biophys Acta. 1221(2):115-24.

Hanauske-Abel HM, Slowinska B, Zagulska S, Wilson RC, Staiano-Coico L, Hanauske AR, McCaffrey T, Szabo P (1995) Detection of a sub-set of polysomal mRNAs associated with modulation of hypusine formation at the G1-S boundary. Proposal of a role for eIF-5A in onset of DNA replication. FEBS Lett. 366: 92–98.

Hanawa-Suetsugu K, Sekine S, Sakai H, Hori-Takemoto C, Terada T, Unzai S, Tame JR, Kuramitsu S, Shirouzu M, Yokoyama S. (2004) Crystal structure of elongation factor P from *Thermus thermophilus* HB8. Proc Natl Acad Sci U S A. 101(26):9595-600.

Hanazawa M, Kawasaki I, Kunitomo H, Gengyo-Ando K, Bennett KL, Mitani S, Iino Y. (2004) The Caenorhabditis elegans eukaryotic initiation factor 5A homologue, IFF-1, is required for germ cell proliferation, gametogenesis and localization of the P-granule component PGL-1. Mech Dev. 121(3):213-24.

Han Z, Sakai N, Bottger LH, Klinke S, Hauber J, Trautwein AX, et al. (2015) Crystal Structure of the Peroxo-diiron (III) Intermediate of Deoxyhypusine Hydroxylase, an Oxygenase Involved in Hypusination. Structure. 23: 882–892.

Hasenöhrl D, Benelli D, Barbazza A, Londei P, Bläsi U. (2006) *Sulfolobus solfataricus* translation initiation factor 1 stimulates translation initiation complex formation. RNA. 12: 674e682.

Hasenöhrl D, Lombo T, Kaberdin V, Londei P, Bläsi U. (2008) Translation initiation factor a/eIF2 (-gamma) counteracts 5’ to 3’ mRNA decay in the archaeon *Sulfolobus solfataricus*. Proc. Natl. Acad. Sci. U.S.A. 105: 2146e2150.

Hoque M, Hanauske-Abel HM, Palumbo P, Saxena D, D'Alliessi Gandolfi D, Park MH, Pe'ery T, Mathews MB. (2009) Inhibition of HIV-1 gene expression by Ciclopirox and Deferiprone, drugs that prevent hypusination of eukaryotic initiation factor 5A. Retrovirology. 6:90

Jakus J, Wolff EC, Park MH, Folk JE. (1993) Features of the spermidine-binding site of deoxyhypusine synthase as derived from inhibition studies. Effective inhibition by bis- and mono- guanylated diamines and polyamines. J. Biol. Chem. 268: 13151–13159.

Jansson BP, Malandrin L, Johansson HE. (2000) Cell cycle arrest in archaea by the hypusination inhibitor N(1)-guanyl-1,7-diaminoheptane. J Bacteriol. 182: 1158–1161.

Jao DL, Chen KY. (2006) Tandem affinity purification revealed the hypusine-dependent binding of eukaryotic initiation factor 5A to the translating 80S ribosomal complex. J Cell Biochem. 97(3):583-98.

Jenkins ZA, Hååg PG, Johansson HE. (2001) Human eIF5A2 on chromosome 3q25-q27 is a phylogenetically conserved vertebrate variant of eukaryotic translation initiation factor 5A with tissue-specific expression. Genomics 71(1):101-9.

Jiang W, Hou Y, Inouye M. (1997) CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. J Biol Chem. 272(1):196-202

Kaiser A. (2012) Translational control of eIF5A in various diseases. Amino Acids. 42(2-3):679-84

Kang HA, Hershey JW. (1994) Effect of initiation factor eIF-5A depletion on protein synthesis and proliferation of Saccharomyces cerevisiae. J Biol Chem. 269(6):3934-40.

Kemper WM, Berry IW, and Merrick WC. (1976) Purification and properties of rabbit reticulocyte protein synthesis initiation factors M2Ba and M2Bß J Biol Chem. 251: 5551–5557**.**

Kim KK, Hung LW, Yokota H, Kim R, Kim SH. (1998) Crystal structures of eukaryotic translation factor 5A from *Methanococcus jannaschii* at 1.8 Å resolution. Proc Natl Acad Sci USA 10419-10424

Klier H, Wohl T, Eckerskorn C, Magdolen V, Lottspeich F. (1993) Determination and mutational analysis of the phosphorylation site in the hypusine-containing protein Hyp2p. FEBS Lett. 334: 360–4.

Laemmli UK. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, pp. 680–685.

Lassak J, Keilhauer EC, Fürst M, Wuichet K, Gödeke J, et al. (2015) Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat. Chem. Biol. 11:266–70. Corrigendum. Nat. Chem. Biol. 11: 299.

Lassak J, Wilson DN, Jung K. (2016) Stall no more at polyproline stretches with the translation elongation factors EF-P and IF-5A. Mol Microbiol. 99(2):219-35

Lecompte O, Ripp R, Thierry J-C, Moras D, Poch O. (2002) Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acids Research. 30 (24):5382-5390.

Lee YB, Joe YA, Wolff EC, Dimitriadis EK, Park MH. (1999) Complex formation between deoxyhypusine synthase and its protein substrate, the eukaryotic translation initiation factor 5A (eIF5A) precursor. Biochem J. 340

Lee NP, Tsang FH, Shek FH, Mao M, Dai H, Zhang C, Dong S, Guan XY, Poon RT, Luk JM (2010) Prognostic significance and therapeutic potential of eukaryotic translation initiation factor 5A (eIF5A) in hepatocellular carcinoma. Int J Cancer. 15: 968–976.

Liao DI, Wolff EC, Park MH, and Davies DR. (1998) Crystal structure of the NAD complex of human deoxyhypusine synthase: an enzyme with a ball-and-chain mechanism for blocking the active site. Structure 6:23–32.

Liu YP, Nemeroff M, Yan YP, Chen KY. (1997) Interaction of Eukaryotic Initiation Factor 5A with the Human Immunodeficiency Virus Type 1 Rev Response Element RNA and U6 snRNA Requires Deoxyhypusine or Hypusine Modification. Biol. Signals 6:166–174

Londei P. (2005) Evolution of translational initiation: new insights from the archaea. FEMS Microbiology Reviews 29 (2005) 185–200.

Londei P. (2007) Translation. USA: ASM Press.

Londei P, Altamura S, Cammarano P, Petrucci L. (1986) Differential features of ribosomes and of poly(U)-programmed cell-free systems derived from sulphur-dependent archaebacterial species. Eur. J. Biochem.157: 455–462.

Lyu Z and Whitman WB. (2017) Evolution of the archaeal and mammalian information processing systems: towards an archaeal model for human disease. Cell.Mol. Life Sci. 74:183-212

Maeda I, Kohara Y, Yamamoto M, Sugimoto A. (2001) Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. Curr Biol. 2001 Feb 6;11(3):171-6.

Mair A, Pedrotti L, Wurzinger B, Anrather D, Simeunovic A, Weiste C, Valerio C, Dietrich K, Kirchler T, Nägele T (2015) SnRK1-triggered switch of bZIP63 dimerization mediates the low-energy response in plants. eLife 4:05828

Mandal A, Mandal S, Park MH. (2014) Genome-wide analyses and functional classification of proline repeat-rich proteins: potential role of eIF5A in eukaryotic evolution. PLoS One. 9: e111800.

Melnikov S, Mailliot J, Shin BS, Rigger L, Yusupova G, Micura R, Dever TE, Yusupov M. (2016) Crystal Structure of Hypusine-Containing Translation Factor eIF5A Bound to a Rotated Eukaryotic Ribosome. J Mol Biol. 428(18):3570-3576.

Nakanishi S, Cleveland JL. (2016) Targeting the polyamine-hypusine circuit for the prevention and treatment of cancer. Amino Acids. 48(10):2353-62.

Navarre WW, Zou SB, Roy H, Xie JL, Savchenko A, Singer A, Edvokimova E, Prost LR, Kumar R, Ibba M, et al. (2010) PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. Mol Cell. 39: 209–221.

Olsen ME, Connor JH. (2017) Hypusination of eIF5A as a Target for Antiviral Therapy. DNA Cell Biol. 36(3):198-201

Park MH. (2006) The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). J Biochem. 139(2):161-9.

Park JH, Dias CA, Lee SB, Valentini SR, Sokabe M, Fraser CS, Park MH. (2011) Production of active recombinant eIF5A: reconstitution in *E. coli* of eukaryotic hypusine modification of eIF5A by its coexpression with modifying enzymes. Protein Eng Des Sel. 24: 301–9.

Park JH, Wolff EC, Folk JE, Park MH. (2003) Reversal of the deoxyhypusine synthesis reaction. Generation of spermidine or homospermidine from deoxyhypusine by deoxyhypusine synthase. J Biol Chem. 278(35):32683-91.

Park MH, Nishimura K, Zanelli CF, Valentini SR. (2010) Functional significance of eIF5A and its hypusine modification in eukaryotes. Amino Acids. 38:491–500.

Patel PH, Costa-Mattioli M, Schulze KL, Bellen HJ. (2009) The *Drosophila* deoxyhypusine hydroxylase homologue nero and its target eIF5A are required for cell growth and the regulation of autophagy. J Cell Biol. 185(7):1181-94

Pedullà N, Palermo R, Hasenöhrl D, Bläsi U, Cammarano P, Londei P. (2005) The archaeal eIF2 homologue: functional properties of an ancient translation initiation factor. Nucleic Acids Res. 33: 1804e1812.

Peil L, Starosta AL, Virumäe K, Atkinson GC, Tenson T, Remme J, Wilson DN. (2012) Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. Nat Chem Biol. 8: 695–697.

Prunetti L, Graf M, Blaby IK, Peil L, Makkay AM, Starosta AL et al. (2016) Deciphering the translation initiation factor 5A modification pathway in halophilic archaea. Archaea. 1–14.

Quintas-Granados LI, Carvajal-Gamez BI, Villalpando JL, Ortega-Lopez J, Arroyo R, Azuara-Liceaga E, Álvarez-Sánchez ME. (2015) Bifunctional activity of deoxyhypusine synthase/hydroxylase from *Trichomonas vaginalis*, Biochimie. 24

Rappsilber J, Mann M, Ishihama Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2:1896–1906

Rossi D, Barbosa NM, Galvão FC, Boldrin PE, Hershey JW, Zanelli CF, Fraser CS, Valentini SR. (2016) Evidence for a Negative Cooperativity between eIF5A and eEF2 on Binding to the Ribosome. PLoS One.11 (4)

Saini P, Eyler DE, Green R, Dever TE. (2009) Hypusine-containing protein eIF5A promotes translation elongation. Nature. 459: 118–21.

Schmidt C, Becker T, Heuer A, Braunger K, Shanmuganathan V, Pech M, Berninghausen O, Wilson DN, and Beckmann R. (2016) Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic Acids Res. 44: 1944–1951.

Schnier J, Schwelberger HG, Smit-McBride Z, Kang HA, Hershey JW. (1991) Translation initiation factor 5A and its hypusine modification are essential for cell viability in the yeast *Saccharomyces cerevisiae*. Mol Cell Biol. 1991 Jun;11(6):3105-14**.**

Shiba T, Mizote H, Kaneko T, Nakajima T, Kakimoto Y. (1971) Hypusine, a new amino acid occurring in bovine brain. Isolation and structural determination. Biochim. Biophys. Acta. 244: 523–531.

Schuller AP, Wu CC, Dever TE, Buskirk AR, Green R. (2017) eIF5A Functions Globally in Translation Elongation and Termination. In: Molecular Cell. 162: 872–884.

Schümann H, and Klink F. (1990) Archaebacterial protein contains hypusine, a unique amino acid characteristic for eukaryotic translation factor 4D. Syst. Appl. Microbiol. 11: 103–107.

Spinozzi F, Beltramini M. (2012) QUAFIT: a novel method for the quaternary structure determination from small-angle scattering data. Biophys J. 2012 Aug 8;103(3):511-21

Spradling AC, Stern D, Beaton A, Rhem EJ, Laverty T, Mozden N, Misra S, Rubin GM. (1999) The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. Genetics. 153(1):135-77.

Tang TH, Polacek N, Zywicki M, Huber H, Brugger K, Garrett R, Bachellerie JP, Hüttenhofer A. (2005) Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. Mol. Microbiol. 55, 469–481.

Turpaev KT. (2018) Translation Factor eIF5A, Modification with Hypusine and Role in Regulation of Gene Expression. eIF5A as a Target for Pharmacological Interventions. Biochemistry (Mosc). 83(8):863-87

Wagner S, Klug G. (2007) An archaeal protein with homology to the eukaryotic translation initiation factor 5A shows ribonucleolytic activity. J. Biol. Chem. 282: 13966–13976.

Wolff EC, Kang KR, Kim YS, Park MH. (2007) Posttranslational synthesis of hypusine: evolutionary progression and specificity of the hypusine modification. Amino Acids. 33(2):341-50

Xu A, Chen KY. (2001) Hypusine is required for a sequence-specific interaction of eukaryotic initiation factor 5A with postsystematic evolution of ligands by exponential enrichment RNA. J Biol Chem. 276(4):2555-61

Xu A, Jao DL, Chen KY. (2004) Identification of mRNA that binds to eukaryotic initiation factor 5A by affinity co-purification and differential display. J Biochem. 384:585-590.

Yanagisawa T, Sumida T, Ishii R, Takemoto C, Yokoyama S. (2010) A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. Nat Struct Mol Biol. 17(9):1136-4

Yao M, Ohsawa A, Kikukawa S, Tanaka I, Kimura M. (2003) Crystal structure of hyperthermophilic archaeal initiation factor 5A: a homologue of eukaryotic initiation factor 5A (eIF-5A) J Biochem. 133(1): 75–81.

Zanelli CF, Maragno AL, Gregio AP, Komili S, Pandolfi JR, Mestriner CA, Lustri WR, Valentini SR. (2006) eIF5A binds to translational machinery components and affects translation in yeast. Biochem Biophys Res Commun. 348(4):1358-66

Zanelli CF, Valentini SR. (2005) Pkc1 acts through Zds1 and Gic1 to suppress growth and cell polarity defects of a yeast eIF5A mutant. Genetics.171(4):1571-81.

Zhang C, Phillips A, Wipfler R, Olsen G, Whitaker R. (2018) The essential genome of the crenarchaeal model *Sulfolobus islandicus.* Nature Commun. 9:4908.

Zuk D, Jacobson A. (1998) A single amino acid substitution in yeast eIF-5A results in mRNA stabilization. EMBO J. 17:2914–2925.