“Sapienza” University of Rome

Department of Experimental Medicine

PhD program

“HUMAN BIOLOGY AND MEDICAL GENETICS”
XXXI cycle

Protein synthesis in Archaea.

1. **Mechanism of dissociation from ribosomes of the conserved translation factor aIF6**

2. **Development of an *in vitro* transcription/translation system from the thermophilic archaeon *Sulfolobus solfataricus*.**

Candidate:
Dr. Giada Lo Gullo

Academic year 2017-2018

Director of doctoral program:

Prof. Antonio Pizzuti

“Sapienza” University of Rome

Scientific Tutor:

Prof.ssa Paola Londei

“Sapienza” University of Rome

Board of examiners:

Prof.ssa Federica Sangiuolo

Tor Vergata University of Rome

Prof.ssa Anna La Teana

Polytechnic University of the Marche, Ancona

Prof.ssa Annalisa Botta

Tor Vergata University of Rome

INDEX

ABSTRACT

This work describes two experimental studies performed for my doctoral thesis. The general subject is the analysis of the features and mechanisms of protein synthesis in extremely thermophilic Archaea. The subject is interesting and relevant from different points of view. Firstly, it has long been known that archaeal translation shares several features and components with the corresponding eukaryal process. Investigating the common themes between archaeal and eukaryal protein synthesis may help to shed light on the evolutionary origin thereof and to reconstruct the pathways whereby Eukarya emerged from the common tree of life. Secondly, extremely thermophilic Archaea have a lot of potential for biotechnological development, for instance as a source of thermostable enzymes to be used for both medical and industrial purposes.

In the first part of my work, I shall describe a functional analysis of the archaeal translation factor aIF6, a protein shared by the Archaea and the Eukarya but not found in Bacteria. Although the eukaryal factor (eIF6) has been studied extensively, its function in translation is still imperfectly understood. It is established that it acts as a ribosome anti-association factor, binding to the large subunit and preventing its premature joining with the small subunit. To allow the 60S ribosome to enter the translation cycle, eIF6 must be actively released, a process that depends on the action of another factor, Sdo1/SBDS, which is also found in Archaea, and of a GTPase, EFL1, that instead does not have an archaeal homologue. In my work, I studied the mechanism of aIF6 release from archaeal large ribosomal subunits, using as the model organism the thermophilic archaeon *Sulfolobus*

solfataricus. The results I obtained show that detachment of aIF6 from ribosomes requires the GTPase activity of elongation factor 2 (aEF2), while the Sdo1/SBDS homologue apparently inhibits rather than promoting aIF6 release. The function of archaeal Sdo1 remains therefore to be understood and must be studied further.

In the second part of work, I developed a coupled *in vitro* transcription/translation system for cell-free protein synthesis from cell lysates of *S. solfataricus*. The essential element in this expression system is a strong promoter derivedfrom 16S/23S rRNA-encoding DNA promoter from the archaebacterium *Sulfolobus sp.* P2 thatproduces, with high efficiency, specific mRNAs. I show thatthis method permits the efficient synthesis *in vitro* at high temperature of biologically active proteins.

1. GENERAL INTRODUCTION

1.1 Archaea

The Domain Archaea was not recognized as a major domain of life until quite recently. Until the 20th century, most biologists considered all organisms to be classifiable as either a plant or an animal. But in the 1950s and 1960s, most biologists came to the realization that this system failed to accommodate the fungi, protists and bacteria. By the 1970s, a system of Five Kingdoms had come to be accepted as the model by which all living things could be classified. At a more fundamental level, a distinction was made between the prokaryotic bacteria and the four eukaryotic kingdoms (plants, animals, fungi and protists). The distinction recognizes the common traits that eukaryotic organisms share, such as nuclei, cytoskeletons, and internal membranes. But the eukaryote-prokaryote concept is fundamentally cytological, and only secondarily phylogenetic. The presumption that the eukaryotic form of cellular organization defines a meaningful phylogenetic unit is a reasonable one. The same is unfortunately not true of prokaryotes, which are united as a class by their lack of the characteristics that define the eukaryotic cell. Thanks to the sequencing revolution, by making accessible the vast store of historical information contained in molecular sequences, the basis for systematic has changed and the classical phenotypic criteria are being replaced by molecular criteria. In the late 1970s, by means of molecular analysis, the scientific community remained surprised by the discovery of an entirely new group of organisms: Archaea. Carl Woese and his colleague George E. Fox, at the University of Illinois, were studying the relationships among the prokaryotes using ribosomal RNAs (rRNAs), the essential components of the protein synthetic machinery that are

encoded by all prokaryotic genomes and in characteristic distinct versions, by eukaryotic nuclear, mitochondrial and chloroplast genomes. They found that the prokaryotic domain, far from being unitary, was split into two distinctly different groups (Woese and Fox,1977): certain micro-organisms that lived at high temperatures or produced methane clustered together as a group well away from the common bacteria and from the eukaryotes as well. Because of this vast difference in genetic makeup, Woese proposed that life be divided into three main domains: Eukaryota, Eubacteria, and Archaebacteria. He later decided that the term Archaebacteria was a misnomer, and renamed it to Archaea. At the cytological level, Archaea are indeed prokaryotes (they show none of the defining eukaryotic characteristics), but on the molecular level (ribosomal proteins, translation factors, RNA polymerase, etc.) they resemble other prokaryotes, the eubacteria, no more than they do the eukaryotes.

A "universal tree of life" based on comparison of rRNA sequences is shown in Fig. 1. Although the studies involving rRNA sequences produce the topology of the universal tree, they do not indicate the position of its root. However, it is in principle possible to root the universal tree using pairs of related (paralogous) genes with distinct functions (Schwarz and Dayhoff, 1978) whose common ancestor duplicated in the ancestral lineage before the three primary lines of descent came into being (e. g. genes coding for translation elongation factors, ATPase subunits, aminoacyl-tRNA synthetases, components of the protein-targeting machinery). Different phylogenetic investigations give the same results: the root lies between the bacteria and the other

two primary lines of descent, which are, thus, sister domains sharing a last common ancestor (Fig. 1).

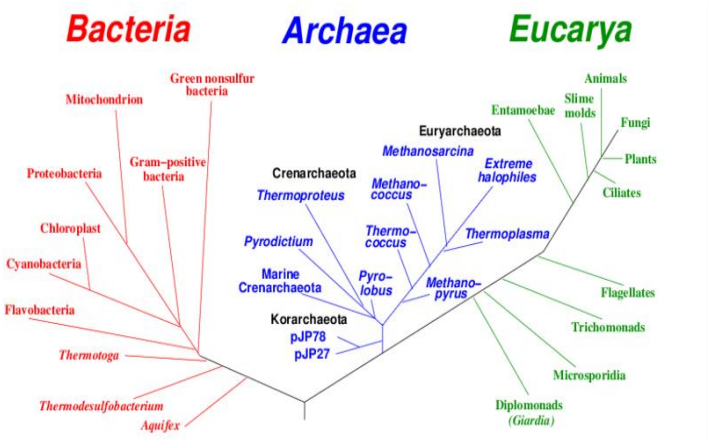

Fig. 1 Rooted Universal Phylogenetic Tree of Life. (Adapted from Woese, 1990).

That the Archaea and Eucarya have indeed shareda common evolutionary path has been proven by the recent analyses of complete genomic sequences. These have shown that the Archaea and the Eucarya share sequence homologies in many genes, especially those encoding components of the gene expression machinery. One example of this evolutionary proximity is given by the transcription apparatus. The RNA polymerases of Eucarya and Archaea resemble each other in subunit composition and sequence far more than either resembles the bacterial type of polymerase. Moreover, whereas all bacterial cells use sigma factors to regulate the initiation of transcription, eukaryal and archaeal cells use TATA-

binding proteins (Langer D. *et al.*, 1995). Other examples ofsimilarities occur in the initiation of archaeal protein synthesis, which makes use of non-formylated methionyl-tRNA; in addition, many archaeal tRNA genes resemble their eukaryal homologs in harbouring introns. Moreover, all Archaea have genes encoding homologues of most eukaryotic translation initiation factors (Kyrpides and Woese, 1998).

1.1.1 Classification

Archaea include inhabitants of some of the most extreme environments on the planet. Some of them live near rift vents in the deep sea at temperatures over 100 °C, others in hot springs or in extremely alkaline or acid waters. They have been found thriving inside the digestive tracts of cows, termites, and the marine life where they produce methane. They live in the anoxic muds of marshes and at the bottom of the ocean, and even thrive in petroleum deposits deep underground. In spite of these environmental differences, Archaea represents a phylogenetically coherent group. From a phenotypic point of view, Archaea are known to comprise four quite distinct general groups: the methanogens, the extreme halophiles, a loosely defined thermophilic ("sulphur-dependent") type, and thermophilic sulphate reducers. However, the four major phenotypes do not correspond to four distinct taxa of equivalent rank. According to the rRNA tree, there are two main branches, corresponding to the kingdoms of *Crenarchaeota* and *Euryarchaeota* (Woese *et al*., 1990). The formers are relatively homogeneous phenotypically, consisting of hyperthermophiles or thermoacidophiles (*Sulfolobus, Desulfurococcus, Pyrodictium, Thermoproteus and*

Thermofilum. The *Euryarchaeota*, on the other hand, constitute an heterogeneous kingdom including hyperthermophiles (i.e. *Pyrococcus* and *Thermococcus*), methanogens (such as *Methanosarcina*), halophiles(some genera are *Halobacterium* and *Haloferax*) and, last but not least, thermophilic methanogens (some genera are *Methanococcus*, *Methanobacterium* and *Methanotermus*). About twenty years ago, PCR analysis of hot spring microbiota detected new archaeal rRNA sequences that branch either deeply within the *Crenarchaeota* or just below the *Crenarchaeota-Euryarchaeota* divergence (Barnes*et al*., 1994). These organisms have been assigned to a third kingdom, the *Korarchaeota* (Barnes*et al.,* 1996). Furthermore, it was found a new nanosized hyperthermophilic archaeon from a submarine hot vent. This archaeon cannot be attached to any of the known groups and therefore must represent an unknown phylum, termed *Nanoarchaeota.*Cells of *Nanoarchaeota equitans* specie are spherical, and only about 400 nm in diameter. They grow attached to the surface of a specific archaeal host and harbor the smallest archaeal genome: only 0.5 megabases in size (Huber *et al*., 2002).

1.2 Archaeal cell biology: genome structure and gene expression

Although the cellular structure of Archaea resembles the bacterial one, consisting of a single cell without both the nuclear membrane and the cytoplasmic organelles, many morphological and metabolic characteristics are distinctive of this kingdom and single the Archaea out of all other organisms. Archaeal membrane lipids are unique in consisting of ether-linkages established between glycerol and hydrocarbon chains, while bacteria and eukaryotes have esters linkages (Kates, 1993). Also, the nature of the layers surrounding the cells differs distinctly from that of bacteria and shows remarkable structural and chemical diversity. The Bacteria (eubacteria) except for the *Chlamydias*, have a semirigid cell wall containing peptidoglycan. Among the Archaea, a cell wall is only found in the order Methanobacteriales; however, it differs from the bacterial one in containing pseudomurein instead of the conventional murein (Kandler, & Koenig, 1993). Moreover, Archaea generally lacking a peptidoglycan cell wall possess surface layers (S-layers) that consist of protein or glycoprotein subunits. The archaeal chromosomes so studied have revealed a single circular double helix of DNA covalently closed (Keeling *et al.*, 1994) whose size is included between 1,7 and 5 Mb, except for *Nanoarcheum equitans*, with a genome of around 0,5 Mb (Huber et al., 2002). Similar to Bacteria lacking a nuclear membrane, Archaea have their genomes organized into a nucleoid, but the architectural proteins that form the nucleoid are largely different. Genomic DNA is packaged into archaeal nucleosomes, which contain an archaeal histone tetramer circumscribed by ~80 bp of DNA. Archaeal histones are not present in *Crenarchaeota* but are present in most *Euryarchaeota*

and resemble in their tetrameric structure the complex formed by the $(H3 + H4)_2$ histone tetramer at the centre of the eukaryal nucleosome (Pereira *et al.*, 1997; Bailey *et al.*, 1999). Similar to Bacteria, and unlike Eukarya, clusters of contiguous genes functionally coordinated tend to be organized in cotranscribed units, producing polycistronic mRNAs. Interestingly, the mosaic nature of archaeal genomes is reflected in the operons themselves, in which there are some genes homologous to genes found in bacteria while other homologous to genes found in eukaryotes (Ouzounis and Kyprides, 1996). Protein-coding genes are not endowed with introns, while these elements have been detected in some RNA-coding genes. For example, in *Sulfolobus solfataricus*, introns have been found in six tRNA (She *et al*, 2001). Archaeal tRNA introns are similar to those of eukaryotes (located one nucleotide 3' to the anticodon (Lykke-Andersen *et al*. 1997; Belfort and Weiner, 1997) but neither group I nor group II self-splicing introns have been reported in Archaea, despite their existence in both Bacteria and Eukaryotes. Moreover, similarly to Eukaryote and Bacteria, Archaea contain inteins, genetic elements present in protein-coding sequences that following translation efficiently remove themselves from the host protein in an autoproteolytic protein-splicing reaction and re-ligate their two flanks. Concerning the replication, many archaeal DNA replication proteins are more similar to those found in Eukarya than Bacteria, including those involved in the initiation process (Grabowski and Kelman, 2003). This suggests that the archaeal origins of replication, their recognition by initiation factors, and the initiation process are also similar to Eukarya. The first archaeal origins were identified using *in silico* analysis, showing a single origin in some Archaea and multiple

origins in others (Grigoriev, 1998; Salzberg *et al*., 1998). The identified origins are rather large, and many are located in close proximity of the genes encoding the archaeal homolog of the eukaryotic initiator protein Cdc6, although the proximity of origin to the Cdc6 gene is not essential for origin function and is not a universal phenomenon in the Archaea domain. *In silico* studies were confirmed by detailed *in vivo* analysis demonstrating the presence of single origins of replication in some species such as *Pyrococcus abyssi* (Myllykallio *et al*., 2000; Matsunaga *et al*., 2001) and *Archeoglobus fulgidus*(Maisnier-Patin *et al*., 2002) and multiple origins of replication in the chromosome of *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius* (Robinson *et al*., 2004; Lundgren *et al*., 2004). Finally, as regards the transcriptional apparatus, Archaea resemble more closely the Eukaryotes than the Bacteria. Archaea do posses a single RNA Polymerase (RNAP) like the Bacteria, but the subunit composition of the archaeal enzyme is far more complex than that of bacterial RNAPs (Keeling and Doolittle, 1995); accordingly, the primary sequence of several archaeal RNAP subunits is closer to that of the eukaryal polypeptides than to that of the bacterial one. The archaeal RNAP, like the eukaryal one, is unable to efficiently recognize promoter sequences on its own and requires two additional basal factors: the TATA-box-binding protein (TBP) and transcription factor B (TFB), homologues of eukaryal TBP and of transcription factor IIB (TFIIB) respectively (Bell and Jackson, 1998). The archaeal promoter resembles in sequence and position the eukaryal RNA polymerase II promoter, which includes a TATA[A/T]A sequence placed approximately 20-30 bases upstream of the transcription start point. In the

Archaea, the TATA box of the eukaryotic promoter is replaced by a "box A" sequence (TTTA[A/T]A) located approximately 27 bp upstream from the transcription start site.

1.3 Protein Synthesis

The degree to which the components of the three major cellular systems (replication, transcription and translation) are conserved in evolution differs greatly. On one side there is genome replication as depicted above, where not even the central DNA polymerase is orthologous between the Bacteria and the Archaea/Eucarya. On the other side there is translation, where many components are well conserved in all domains: ribosomal RNAs, the tRNAs and the elongation factors. Translation by the ribosome is traditionally divided into several separate steps: initiation, elongation, termination and recycling. All of these steps are aided by translation factors, which accordingly, are called initiation, elongation, release and recycling factors. At least ten protein factors participate in bacterial translation, whereas a considerably larger number is needed in eukaryotic organisms (Sachs *et al*., 1997; Dennis, 1997). Several of these factors are GTPases, usually activated after binding to the GTPase centre of ribosome. Translation factors and tRNAs bind to the ribosome transiently and move between different binding sites. The ribosomal tRNA binding sites are the decoding A-site, where amino acyl tRNA binds, the P-site, where the peptidyl tRNA binds, and the E-site, where deacylated tRNA binds prior to its dissociation from the ribosome. During translation initiation, ribosomes identify the initiation codon on the mRNA and set the correct reading frame for decoding. This rate-limiting

step is carried out in different ways in the primary domains of cell descent. In Bacteria, the small ribosomal subunit interacts directly with the mRNA by base-pairing of the Shine-Dalgarno (SD) motif, preceding the initiation codon, with the anti-SD motif at the 3' end of 16S rRNA. The initiation step is assisted by three protein factors, IF1, IF2 and IF3 (Gualerzi and Pon, 1990). In eukarya, selection of the initiation codon entails a "scanning" mechanism, whereby the 40S subunit binds to the capped 5' end of the mRNA and then slides downstream until the initiator AUG, usually the first one available, is encountered (Kozak, 1983). This process is promoted by many factors, several of which are involved in cap recognition and mRNA unwinding, while others interact with the ribosome and/or the mRNA to ensure correct selection of the initiation codon (Pestova and Kolupaeva, 2002).

1.4 Archaeal protein synthesis

In Archaea, the absence of a nuclear membrane does provide the opportunity for ribosomes to bind to a transcript and to initiate translation before the transcript is complete, and this could be a positive feature of being a prokaryote. (Martin and Koonin, 2006). Coupling of transcription and translation has been documented and studied extensively in Bacteria, where it provides the molecular basis for regulation of gene expression by attenuation (Grundy and Henkin, 2006). It seemed reasonable to predict that Archaea would also have exploited the regulatory opportunities offered by coupling transcription and translation. Archaeal genome sequences and northern blot results do argue convincingly that many archaeal genes are co-transcribed as members of multigene operons, but only one putative

attenuator has been identified, and this seems likely to be regulated by a riboswitch rather than by a translating ribosome (Rodionov *et al*, 2003). Furthermore, archaeal genomes do not encode detectable homologues of the transcription termination factor rho (ρ) nor have functionally equivalent termination factors been identified. Furthermore, although archaeal RNA polymerases do exhibit intrinsic termination, they terminate in response to DNA template sequences without an apparent requirement for nascent transcript folding (Santangelo and Reeve, 2006). Evidence for coupled transcription and translation in Archaea was provided by electron microscopic examination of chromatin extruded from archaeal cells of the marine hyperthermophile *Thermococcus kodakaraensis*, which clearly revealed the presence of polysomes containing up to 20 adjacent ribosomes attached directly to dispersed strands of the archaeal genomic DNA. The polysome patterns were consistent with the sequential direct binding of ribosomes to nascent mRNAs. It is noteworthy that, although many archaeal genes are preceded by Shine-Dalgarno sequences functioning as ribosome-binding sites, a large proportion of archaeal trascripts, especially in certain species, have no leader sequence at all or only very short ones (Torarinsson *et al*, 2005).

Initiation of protein synthesis consists of several interrelated events that take place before the formation of the first peptide bond. The main goal of this step is the correct selection of the first codon on the mRNA by means of a specialized initiator tRNA (tRNAi), which becomes adapted in the ribosomal P site, ready to accept the next amino acid. Initiation terminates with the joining of the two ribosomal subunits, which allows the elongation phase to begin. In Archaea, two different

mechanisms for translational initiation exist. One is based on a canonical SD/ anti-SD interaction and operates mostly on internal cistrons of mRNAs. In contrast, monocistronic mRNAs as well as proximal cistrons of polycistronic mRNAs, which are frequently devoid of a 5'-untranslated region, are decoded by an initiation mechanism independent of the SD/anti-SD interaction, which requires pairing of the start codon with initiator-t RNA (tRNAi) (Benelli *et al*, 2003). The complexity of archaeal translational initiation is underscored by the presence of a larger-than-bacterial set of initiation factors. All archaeal genomes encode about 10 proteins homologous to eukaryal initiation factors (Dennis, 1997). One prominent example is the factor that adapts tRNAi in the ribosomal P site. This task is assisted in bacteria by the monomeric protein IF2, while in Eukarya and in Archaea it is performed by the trimeric complex eIF2, consisting of the α-, β-, and γ-subunits, none of which is homologous to bacterial IF2. However, both Eukarya and Archaea possess a homologue of bacterial IF2, termed eIF5B, which seems to act at a later initiation step, promoting the joining of the 60S ribosomal subunit to the 40S initiation complex (Pestova *et al*, 2000). With the exception of a/eIF2, the function of the eukaryal-like translation initiation factors in Archaea is poorly understood.

Chapter 1.

Analysis of the mechanism of aIF6 removal from *S. solfataricus* ribosomes.

Introduction

The conserved translation factor IF6

The Archaea and the Eukarya selectively share several translation factors. One of these is the small monomeric protein (about 25 kDa) called eIF6 in eukarya and aIF6 in Archaea. The function of this factor has been extensively studied in eukaryotes, remaining however somewhat elusive. Much less is known about archaeal IF6.

1.5 IF6 in eukaryotes

The eukaryotic Initiation Factor 6 is a protein of 245 aminoacids and 77% identical between yeasts and humans (Biffo *et al,* 1997). The eIF6 structure is organized in a peculiar cyclic, "star-like" structure named pentein (Fig.2).

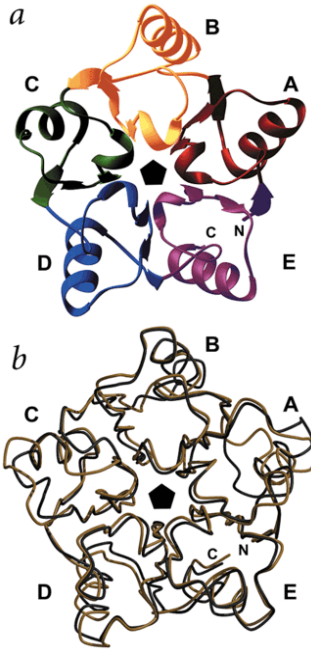

Fig. 2. Pentein structures of aIF6 and eIF6. a), RIBBONS drawing of *M. jannaschii* aIF6 (top), with a, black pentagon denoting the five-fold pseudosymmetry axis. Individual aIF6 subdomains are colour coded in progression from the N-terminus as follows: A, red; B, yellow; C, green; D, blue; E, purple. b), Cα backbone traces of the *M. jannaschii* aIF6 (black) and *S. cerevisiae* eIF6(gold) overlaied to show local differences. (Source: adapted from Groft*et al*. (2000) in *Nature structural biology* 7(12): 1156-1164.

eIF6 is formed by five stretches of quasi-identical alpha/beta-sub domains arrayed around a fivefold axis of pseudo-symmetry. It is a rigid protein, with a cavity that contains 16 well-ordered water molecules with a limited degree of motility. The semi-conserved carboxy-term was proposed as a candidate region for eIF6 regulation due to its flexibility (Groft *et al*, 2000). Accordingly, the carboxy-term of eIF6, in mammalian cells, is characterized by the presence of several phosphorylation sites whose significance is not totally understood. In Archaea, IF6 has a very similar structure, lacking however the C-terminal tail (Fig.2).

eIF6 is able to bind the 60S ribosomal subunits and prevent their joining to the 40S ribosomal subunits.

(Valenzuela *et al*, 1982). The main interaction partner of eIF6 on the 60S, is the ribosomal protein rpL23. The C-terminus of rpL23 mediates the interaction with the major binding surface of eIF6. (Klinge *et al*, 2011). Also, the neighboring ribosomal protein rpL24 and the highly conserved sarcin-ricin loop (SRL) contribute to the interaction of 60S with eIF6. In conclusion, all data converge on demonstrating that eIF6 binding to the 60S impedes 40S recruitment, thus preventing premature 80S formation. Therefore, to promote translation, a regulated mechanism for eIF6 release is required.

Two mechanisms have been proposed for eIF6 release in eukaryotes. One relies on the fact that eIF6 interacts in the cytoplasm with RACK1, a receptor for activated protein kinase C(PKC). External stimuli activate the PKC cascade that leads to eIF6 phosphorylation through RACK1. Phosphorylated eIF6 is released from 60S subunits, with the final result of ribosome activation.

Another described mechanism for eIF6 release is based on the joint action of the Sdo1/SBDS protein and of the elongation-factor-like protein EFL1. The former stimulates 60S-dependent GTP hydrolysis by EFL1, with consequent release of eIF6 and formation of actively translating 80S ribosomes. Specifically, SBDS and EFL1 jointly evict the anti-association factor eIF6 from the intersubunit interface of 60S ribosomal subunits to allow ribosomal subunit joining. Notably, SBDS mutations that uncouple GTP hydrolysis from eIF6 release cause the Shwachman-Diamond syndrome after which the protein is named (Finch *et al*, 2011).

The two proposed mechanisms are not mutually exclusive. They may act in different cell lines or in specific physiological situations.

Notably, eIF6 localizes both in the nucleolus, where it is enriched on the perinucleolar region, and in the cytoplasm (Lebreton *et al*, 2006). The presence of eIF6 in the nucleolus supports a role in ribosome biogenesis. In *Saccharomyces cerevisiae*, the ribosome biogenesis pathway begins with transcription of the 35S and 5S ribosomal RNA (rRNA) precursors by RNA polymerases I and III, respectively. The association of ribosomal proteins and pre-ribosomal factors with nascent pre-rRNA originates a 90S pre-ribosomal complex. The 90S complex separates into a pre-60S complex, which in turn generates the large ribosomal subunit containing mature 25S, 5.8S and 5S rRNAs and a pre-40S complex, which gives rise to the small ribosomal subunit containing 18S rRNA (Lempiainem, 2009). In agreement with its proposed role in ribosome biogenesis, eIF6 is found in large molecular weight complexes of 60S pre-ribosomes (Volta *et al*., 2005). In mammals, however, most of eIF6 is present in the cytoplasm (Sanvito *et al*., 1999).

The above observations have led to propose that 60S subunits carrying eIF6 would be the newly synthesized ones, just shipped to the cytoplasm. eIF6 release would be the last step of subunit maturation, that would act also as a sort of quality control.

1.6 EFL1 and SBDS/Sdo1 function in eIF6 release

The elongation-factor like GTPase EFL1, that is required for the release of eIF6 in yeast, is highly homologous to eEF2. It has been shown that EFL1 can compete with eEF2 for ribosome binding resulting in inhibition of the eEF2's ribosome associated GTPase activity. Therefore, it is safe to assume that they share a similar binding site that is known for eEF2 to be the canonical translation factor binding site.

EFL1 has the basic organization of a translocation factor composed of the G domain (domain I) and domains II–V (Fig. 3). The G domain that binds and hydrolyses GTP consists of five highly conserved motifs (boxes G1–G5). Both the relative position of the boxes within the G domain and the residues known from thecrystal structures to be involved in the interactions with GTP are conserved in EFL1 (Fig.3A) (Berchtold*et al*,1993; Czworkowski *et al*, 1994). A distinguishing feature of EFL1 is the presenceof a 160 aminoacid insertion between the G domainand domain II that has no equivalent among other EF 2-like factors (indicated in yellow in Fig. 3); 40% ofthis insertion consists of small stretches of acidic/serine residues. Deletion of the insertion domain results in athermosensitive phenotype. Homologs of EFL1 are present in *Homo sapien*s, *Arabidopsis thaliana*, *Drosophila melanogaster*, and *Schizosaccharomyces pombe*, indicating that the protein is highly conserved.

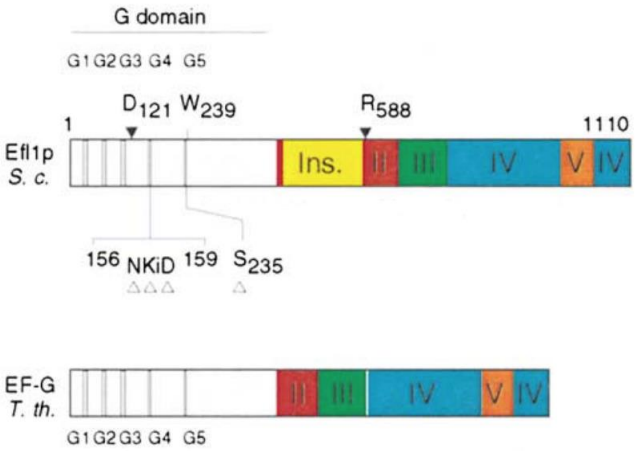

Figure 3. EFL1 is a G protein homologous to the family of EF-2/EF-G translocases. Diagram of the primary structure of EFL1 from *S. cerevisiae*. The five domains defined by the crystal structure of EF-G from *T. thermophilus* are color coded as follows: domain I (G domain in black), domain II (red), domain III (green), domain IV (blue), and domain V (orange). Open arrowheads indicate functional residues from the G4 and G5 boxes that are involved in guanine base recognition. Closed arrowheads indicate conserved amino acids involved in salt bridge formation between domains I and II. The highlighted tryptophan (position 239 in EFL1p) was found to be important for the interaction with the ribosomal RNA sarcin/ricin loop. The 160-amino acid insertion within domain II is indicated in yellow.

Cryo-EM, biochemical and genetic analyses permitted to propose a mechanism for eIF6 release (Fig. 4). The model is based on a cofactor-dependent conformational-switching mechanism in which EFL1 initially binds to the GTPase center, in direct contact with SBDS and eIF6, in a low-affinity, inactive GTP-bound state (Fig. 4a). Competing with SBDS for an overlapping binding site, EFL1 domain V promotes a 180° rotational displacement

of SBDS domain III away from the P-stalk base (closed state) toward helix 69 (open state), causing the SBDS protein to adopt a conformation that is probably stabilized by interactions between SBDS residues K151 and R218 and helix 69(Fig. 4b). In the open state, SBDS drives the equilibrium of GTP-bound EFL1 toward an active high-affinity ('accommodated') SRL-bound conformation that effectively competes with eIF6 for an overlapping binding site on the SRL and promotes eIF6 displacement from the 60S subunit (Fig. 4c). In the final step of the catalytic cycle, the interaction of EFL1 with the SRL promotes GTP hydrolysis, thereby shifting the EFL1 conformational equilibrium from a high- to a low-affinity ribosome binding state and promoting dissociation of both EFL1 and SBDS from the 60S subunit (Fig. 4d).

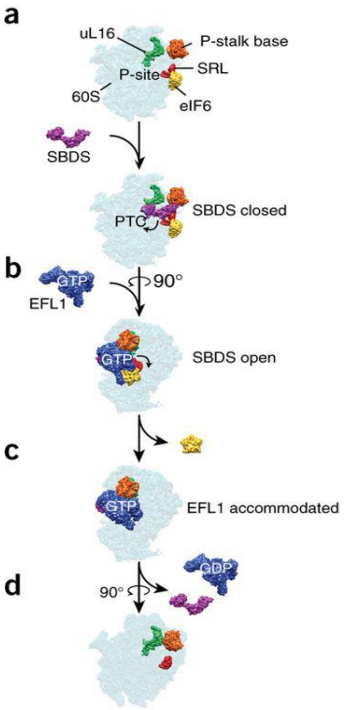

Figure 4. Mechanism of eIF6 release by SBDS and EFL1. (a) SBDS (closed state) is recruited to a late cytoplasmic eIF6-loaded pre-60S subunit **(b)**EFL1–GTP binds directly to SBDS and eIF6 in the GTPase center, thus promoting rotational displacement (180°) of SBDS domain III away from the P-stalk base toward helix 69 (open state), which is stabilized by SBDS residues K151 and R218. **(c)** GTP-bound EFL1 in the accommodated state competes with eIF6 for an overlapping binding site on the 60S subunit, thus promoting eIF6 displacement. **(d)** Interaction of EFL1–GTP with the SRL promotes GTP hydrolysis, thus triggering a conformational switch in EFL1 that promotes a low-affinity ribosome binding state. SBDS and EFL1–GDP dissociate from the 60S subunit.

1.7 IF6 in Archaea

All archaeal genomes possess a horthologue gene of eIF6. The Archaeal and eukaryal IF6 proteins share a considerable degree of homology in their primary sequence and have the same tertiary folding (Fig.2), suggesting that they share a core function conserved in the eukaryal/archaeal line. The function of aIF6 has been much less studied than that of its eukaryal homologue, however the available data suggest a marked functional similarity between the two proteins.

Sulfolobus solfataricus aIF6(aIF6) acts as a translational inhibitor by binding specifically to the large ribosomal subunit and impairing the formation of 70S particles. That aIF6 may function also *in vivo* as a translational repressor under unfavourable conditions is suggested by the fact that it is over-expressed upon both cold-and heatshock (Benelli *et al*, 2009). It is, instead, doubtful whether aIF6 has a function in ribosome synthesis as proposed for its eukaryal counterpart. Firstly, aIF6 is expressed at about the same level in different growth phases of *S. solfataricus* cells, while a ribosome synthesis factor is expected to be up regulated during exponential growth.Secondly, aIF6 is over-expressed upon thermal shock, a circumstance in which most ribosomal genes are down regulated. Thirdly, aIF6 is present in sub-stoichiometric amounts with respect to the 50S subunits (about 1:10), but the aIF6 binding site remains available on the entire cellular pool of large ribosomal particles, which is not expected if aIF6 dissociation is a final step in large subunit maturation.

These data suggest that the protein may have evolved in the archaeal/eukaryal lineage to fulfil a main role in translational regulation. Eukaryal IF6 may then have

acquired additional functions during the evolution of the eukaryotic domain. However, like its eukaryal counterpart, aIF6 is tightly bound to the large ribosomal subunit, and must be released for the subunit to be active in translation. My work addresses the possibile mechanism whereby such release is carried out.

2. MATERIALS AND METHODS

2.1 Preparation of *S. solfataricus* cellular extracts and other cellular fractions

About 5 g of frozen *Sulfolobus solfataricus* cells were disrupted by grinding with 10 g of alumina powder while gradually adding 2.5 ml/g wte weight of cells of ribosome extraction buffer (20 mM Tris-HCl pH 7.4, 10 mM $Mg(OAc)_2$, 40 mM NH_4Cl, 1 mM DTT, 2.5 µg/ml RNase-free DNase). Alumina and cellular debris were removed by centrifugation twice at 30,000x *g* for 30 min. The clarified supernatant obtained **(S-30)** was stored at -80°C and total protein concentration, determined by Bradford assay, was in the range of about 20-25 mg/ml according.

Crude cellular lysates (S30) were centrifuged in a Beckman Ti 50 rotor at 100,000x *g* and 4° C for 3 h to separate ribosomes from a supernatant **(S-100)** containing total cellular tRNAs and ribosome-free cytoplasmic proteins.

Unfractionated tRNA from *S. solfataricus* was prepared performing a phenol extraction of the crude S-100 fraction and precipitating the aqueous phase with 2.5 volumes of 95% ethanol.The RNA pellet was resuspended in 10 mM glycine pH 9.0 and the solution was incubated 2h at 37 °C to achieve alkaline deacylation of the tRNA therein contained. Lastly, the RNA was again precipitated and the resulting pellet was dissolved in an adequate volume of 10 mM Tris-HCl (pH 7.5).

The pellet of **ribosomes** (termed crude ribosomes) was further purified by resuspending in buffer (20 mM Tris-HCl pH 7.4, 500-mM NH_4Cl, 10-mM $Mg(OAc)_2$, 2 mM dithithreitol and applying on 18% (w/v) sucrose gradient in the same buffer to centrifuge in a Beckman Ti 50 rotor at 100,000x *g* for 6 h at 4°C.

The final pellets containing some amount of certain translation factors (e.g., aIF6), was resuspended in extraction buffer. This preparation represents the purified ribosomes. In order to devoid of translation factors ribosomes, these purified ribosomes were resuspended in a high salt buffer. The ribosome suspension was stirred on ice for ~1 h to allow the release of the nonribosomal proteins. The mixture was then layered onto a 7.0-ml cushion of 0.5 M sucrose made in high-salt buffer and centrifugated in a Beckman Ti 50 rotor operated at 45,000 rpm for 3 hours (4°C). The ribosome pellet (termed **"high-salt-washed ribosomes"**) are lastly resuspended in ribosome extraction buffer containing 10%(v/v) glycerol.

The concentration of the ribosomes was determined by measuring the A_{260} and using as the extinction coefficient 1 OD_{260} 70S =40 pmol.

The supernatant recovered after the sedimentation of the high-salt-washed ribosomes was supplemented with 70% (final concentration) ammonium sulphate and placed on ice for ~ 1 h to allow protein precipitation.

The precipitate was collected by centrifuging 10 min at 15,000 rpm; the pellet was dissolved in resuspending buffer (20 mM Tris-HCl pH 7.4,2 mM dithithreitol, 10% glycerol) and dialyzed against the same buffer. This preparation was the high-salt ribosome wash (HSW).

2.2.1 Cloning of aSdo1 gene

In order to clone the aSdo1gene, two synthetic DNA oligomers were constructed on the sequence of *S. solfataricus*(*Sso*0737), the homologous of Sdo1, as deduced from the published genome sequence. The primers permitted us to clone the gene by PCR amplification; the forward aSdo1 is 5'-TTTTTTTATGCTAGCATGACGAAGGAGCGTGATT ATG-3'; the reverse aSdo1 is 5'-CATGGTATGCTCGAGTCATCTCACTTGCAATACT TTAAC-3'. Furthermore, the forward and reverse primers contained the XhoI and NheI sites respectively, which allowed insertion of the gene into the NheI/ XhoI sites of expression plasmid pRSETB (+) (Novagen). The vector adds to the N-terminus of the protein a tag of six histydine residues and the transcription of the cloned gene is directed by the promoter of phage T7. To ascertain the correct cloning of the gene, the purified recombinant vector was sequenced.

2.2.2 Preparation of recombinant aSdo1 under native conditions.

E.coli strains BL21, whose genome carries the RNA polymerase T7 gene under the control of lac UV5 promoter, were transformed with the plasmid vector pRSETB(+)-His_6-aSdo1 and grown at 37°C in LB medium containing ampicillin(100 µg/ml). Expression was induced at OD=0.6 with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) and the cells grown for a further 3 hours before harvesting. The cell pellet was resuspended in Lysis buffer (50 mM NaH_2PO_4, 300 mM

NaCl, 10 mM imidazole, pH 8.0) and incubated with lysozyme on ice for 30 minutes.
The cells were lysed by sonication and the lysate was clarified by centrifugation at 5000 rpm for 20 minutes at 4°C. Thermostable His_6- aSdo1 was purified from the lysate incubating overnight on Ni-NTA agarose resin (Qiagen), washing with Wash buffer (50 mM NaH_2PO_4, 300 m M NaCl, 20 mM imidazole, pH 8.0) and eluting with Elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0).Then, this preparation was dialyzed against Storage buffer (20 mM TEA pH 7.4, 10 mM KCl, 10% glycerol) and the concentration of the samples was determined by the Bradford assay. Aliquots of protein were stored at -80°C.

2.3.1 Cloning of aEF2 gene.

To clone the aEF2 gene, two synthetic DNA oligomers were constructed on the sequence of *S. solfataricus*(*Sso*0728) as deduced from the published genome sequence. The primers permitted us to clone the gene by PCR amplification; the forward aEF2 is 5'-TTTTTCCATGGCTTGCCTAGATATAAGACAGTAG AGC -3'; the reverse aEF2 is 5'-TTTTTGGATCCTCACGACAAGAAATCTTCCACTT TTGGC -3'. Furthermore, the forward and reverse primers contained BamHI and NcoI sites, respectively, which allowed insertion of the gene into the NcoI/BamHI sites in correspondence of the MCS carried by the expression plasmid pETM11(+) (Novagen). Specifically, this vector adds a tag of six histydine residuesto the C-terminus of the recombinant proteinand

the transcription of the cloned gene is directed by the promoter of phage T7. To ascertain the correct cloning of the gene, the purified construct was sequenced.

2.3.2 Preparation of recombinant aEF2 under native conditions.

E.coli strains BL21, whose genome carries the RNA polymerase T7 gene under the control of lac UV5 promoter, were transformed with the plasmid vector pETM11(+)-His_6- aEF2 and grown at 37°C in LB medium containing kanamycin (30 µg/ml). Expression was induced at OD=0.6 with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) and the cells grown for a further 3 hours before harvesting. The cell pellet was resuspended in Lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) and incubated with lysozyme on ice for 30 minutes.

The cells were lysed by sonication and the lysate was clarified by centrifugation at 13000 *g* for 20 minutes at 4°C. Thermostable His_6- aEF2 was purified from the lysate incubating overnight on Ni-NTA agarose resin (Qiagen), washing with Wash buffer (50 mM NaH_2PO_4, 300 m M NaCl, 20 mM imidazole, pH 8.0) and eluting with Elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0).

Then, this preparation was precipitated by adding $NH_4(SO_4)_2$ and dialyzed against Storage buffer (20 mM Tris/HCl pH 7.4, 10 mM KCl, 10% glycerol. The pellet was resuspended in a suitable dialysis buffer (30 mM NH_4Cl, 20 mM Tris/HCl, pH 8.0, 5% glycerol; finally, the concentration of the samples was determined by the Bradford assay. Aliquots of protein were stored at -80°C.

2.4 Polysomal profiles

Ribosomes from *S. solfataricus* were layered on top of a linear 10-30% (w/v) sucrose gradient containing 1M Tris-HCl pH 6.8, 1M KCl, 1M $Mg(OAc)_2$. Reactions were incubated at 65°C for 10 min. After the gradients were centrifuged at 4°C in a SW41 Beckman rotor for 18,000 rpm overnight and unloaded. Fractions (0,5 ml) were collected in 18 tubes and precipitaded with TCA-DOC. To one volume of protein solution add 1/100 volume of 2% DOC (Na deoxycholate detergent). We vortex the sample and let sit for 30 minutes at 4°C. We add 1/10 of Trichloroacetic acid (TCA) 100%, vortex and let sit over night at 4°C. Subsequently we spin at 15000*g* for 15 minutes at 4°C, carefully discard supernatant and retain the pellet. Wash the pellet twice with cold 100% of acetone. At this point we do dry the pellet and we prepare the samples for SDS-PAGE resuspending in a minimal volume of sample buffer.

2.5 Western blotting analysis

The protein levels were analyzed by Western blot. Before electrophoretic analysis, the total protein concentration in cell lisates was determined by Bradford assay. Then, each recombinant protein was loaded on SDS-polyacrylamide gels under reducing conditions. After electrophoresis, the proteins were electroblotted onto nitrocellulose membranes (Amersham Protran); then, the membranes were incubated overnight with primary antibody after blocking. Membranes were developed using the enhanced chemiluminescence method (ECL Western Blotting Substrate, Pierce). The bands were quantified by densitometric analyses using Image Lab program (Bio-rad) and the total amount of proteins were visualized by Comassie Blue R-250 staining. To reveal both the aSdo1 and the aEF2 proteins, the primary antibody used was Penta-His-antibody (Qiagen).

2.6 GTPase assay

It is an assay of inorganic phosphate in which the phosphomolybdate complex is reduced by ascorbic acid consisting in a mix of Ammonium Molybdate-Sulfuric Acid solution,5%. The reagents used for determination of phosphate were: 1 part of ascorbic acid,10% to 6 parts of 0,42% Ammonium Molybdate.$4H_2O$ in 1 N H_2SO_4. We add 0,70 ml of the mix (called also Ames reagent) to 0,30 ml of the phosphate solution and incubate 20 minutes at 45°C. Before the incubation, we prepared aliquots in a reaction volume of 50 µl in the presence of 100 mM Tris/HCl pH 7.4, 100 mM $MgCl_2$, 100 mM KCl, 100 mM GTP. The readings at 660 nm(UV) are proportional to

phosphate concentrations to an optical density of at least 1.8. We monitored the amount of inorganic phosphate released after the GTP hydrolysis at 65°C at different times.

2.7 Size-exclusion chromatography(SEC)

The size-exclusion or gel filtration chromatography is a technique in which molecules in solution are separated by their size, and in some cases molecular weight.An aqueous solution is used to transport the sample through the column. The chromatography column is packed with fine, porous beads, which are composed of dextran polymers (Sephadex). The pore sizes of these beads allows to estimate the dimensions of macromolecules of our interest.Small molecules diffuse freely into the pores and their movement through the column is retarded, whereas large molecules are unable to enter the pores and are therefore eluted earlier. Hence, molecules are separated in order of decreasing molecular weight, with the largest molecules eluting from the column first.Allyl dextran-based size-exclusion gel (Sephacryl S-300) was used as the gel media. A gel column was prepared by filling the same amount of allyl dextran-based size-exclusion gel (1 ml of gel beads) into thecolumn. After the gel column was equilibrated with a solution, aliquots of the mix reaction were applied to the colomn. Next, a solution was used to elute the unbound in the eluent. After, fractions (0,5 ml) were collected in 18 tubes and precipitated with TCA-DOC. To one volume of protein solution add 1/100 volume of 2% DOC (Na deoxycholate detergent). The samples were vortexed and let sit for 30 minutes at 4°C. Then, 1/10 of

Trichloroacetic acid (TCA) 100%, was added and let sit over night at 4°C. Subsequently, we spun at 15.000*g* for 15 minutes at 4°C.The retained pellets were washed twice with cold 100% of acetone, dried and resuspended in an adeguate volume of Laemelly buffer 1x. The analysis of the precipitated proteins was performed by 15% SDS-PAGE gels and detecting their presence by Western analysis.

3. RESULTS

3.1 70S particles are affected by the presence of aIF6

The anti-association and translational inhibitory action of *S.solfataricus* aIF6 is illustrated in Fig. 1. The graphs show that when recombinant aIF6 is added in excess to a cell lysate programmed for *in vitro* translation, the amount of 70S monomers in the samples decreased with increasing amounts of added protein (Benelli *et al*., 2009).

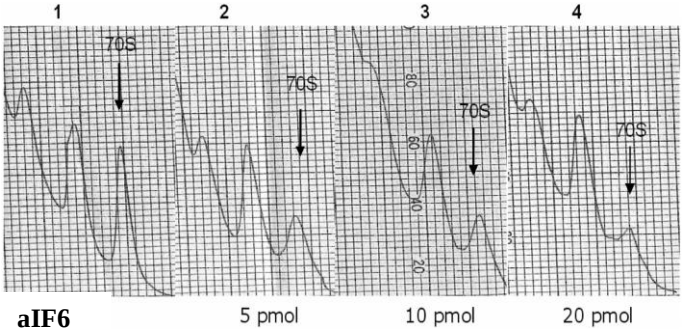

Fig. 1. Excess aIF6 impairs 70S formation. 20 µl of lysate were incubated with increasing amounts of aIF6 at 73°C for 15 minutes. At the end of reaction, the samples were loaded on sucrose density gradient and, after centrifugation, the ribosome profile was checked.

In *Sulfolobus*, most of cellular aIF6 is in a 60S-ribosome-bound state. The aIF6-ribosome interaction is very strong, as shown by the fact that the factor cannot be removed with high-salt washing procedures routinely used to free the ribosomes of extrinsic, loosely-bound proteins (Benelli *et al*, 2009). We presume, therefore, that aIF6 must be released by some specific mechanism as it happens in the eukaryotes, and that this mechanism would serve to regulate the access of large ribosomal subunits to the elongation cycle.

3.2 Cloning of *S.solfataricus* aSdo1

It is known that all archaeal genomes contain a homologue of eukaryotic SBDS, the factor required for eIF6 release from ribosomes. Therefore, we wished to learn whether the archaeal SBDS homologue (termed aSdo1 here) had a similar function. To obtain some insight about this problem, we decided to clone the *S. solfataricus* aSdo1 gene and to obtain the relative purified recombinant protein using conventional overexpression/purification systems. On the basis of the genome sequence of *S. solfataricus*, a couple of PCR primers were designed to amplify the aSdo1 gene by PCR. The product of the reaction was inserted into an expression vector (pRSETb+) that added a ($His_{6)}$-tag to the C-terminus of the protein, whose transcription was directed by the T7 promoter. The construct was then expressed in *E. coli* BL21(DE3) strain, whose genome carries the RNA polymerase T7 gene under the control of a lac UV5 promoter. Gene expression was induced by IPTG; after 3 h of induction, the amount of expressed protein was sufficient for subsequent purification. The

aSdo1protein was purified from *E. coli* extracts under native conditions and through affinity chromatography on a Ni^{2+}agarose column that binds selectively the His_6-tagged protein. (Fig. 2).

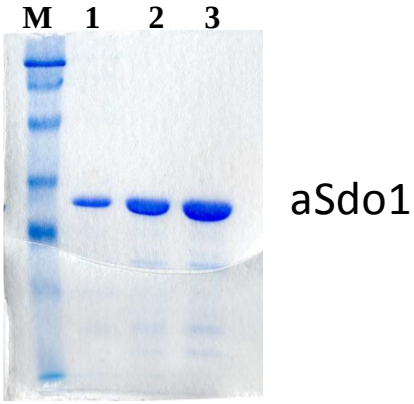

Fig. 2. Expression and purification of aSdo1gene was cloned in the pRSETb+ expression plasmid and the construct was expressed in *E.coli* BL21(DE3) strain. At the end of expression, the recombinant protein was purified by Ni^{2+}agarose columns. The product of purification was loaded on 15% SDS-PAGE and stained with Coomassie **M**, molecular weight marker; **1-3)** Increased amount of recombinant aSdo1corresponding to 50, 100 and 150 pmol, respectively.The arrow indicates the recombinant protein aSdo1.

The recombinant aSdo1 was firstly used to prepare specific antibodies, which were employed to determine the cellular distribution of the protein. To this end, the *S.solfataricus* S-30 extract was fractionated on sucrose density gradient and the position of the factor with

respect to the ribosomes was monitored by Western blot with anti-His antibodies. As illustrated in Fig.3, the endogenous aSdo1 protein was associated with both 30S and 50 S ribosomes, but most of it was localized in the pre-ribosomal fraction.

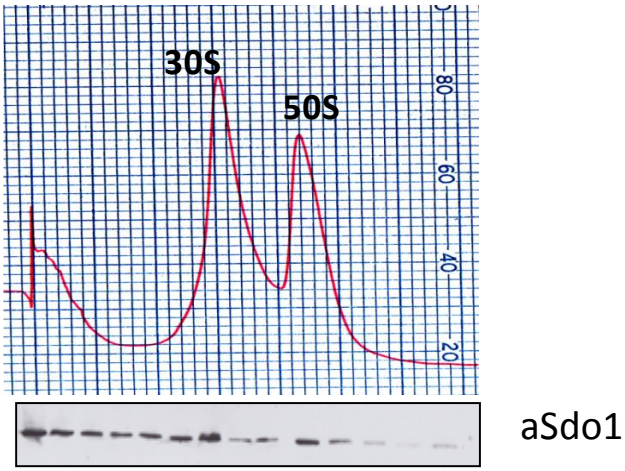

Fig. 3. Sucrose density gradient of *S. solfataricus* extract incubated with aSdo1. The reaction was layered on linear 10 to 30% sucrose gradients. Upper, profile of ribosome distribution along the sucrose gradient. The gradient fractions were checked for both absorbance at 260 and the presence of aSdo1 with the specific antibodies (lower panel).

Next, to determine the strength of the aSdo1-ribosome interaction, "crude" *S.solfataricus* 70S ribosomes were treated with 0.5 M NH_4Cl (salt-wash), a procedure routinely employed to remove from the ribosomes any extrinsic, loosely associated proteins. We note that, as said above, high-salt washing does not dissociate aIF6

from the ribosomes. After salt-washing, the different fractions (salt-washed ribosomes and their supernatant fraction, hereinafter termed HSW) were run on gels and subjected to western blotting.
The results of these experiments showed that the protein aSdo1 is present on crude 70S ribosomes but it is detached from the ribosomes when they are washed with high salt concentrations. (Fig. 4). We concluded that, unlike aIF6, aSdo1 is loosely associated with the ribosomes.

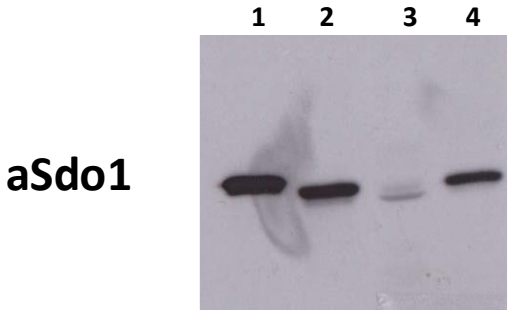

Fig. 4. Western blotting for aSdo1. Lane 1, the recombinant aSdo1; lane 2, aSdo1 band resulted in HSW; lane 3, no band correspond to protein in high-salt washed ribosomes (70SHSW); lane 4, the aSdo1 detected in 70S crude. All *S. solfataricus* fractions were loaded on 15% SDS-PAGE before electroblotting onto nitrocellulose membrane.

The presence of aSdo1 on both 30S and 50S subunits was somewhat puzzling, as it is known that eukaryotic SBDS resides specifically on the 60S subunit. Since, as shown before, aSdo1 is not strongly associated with the ribosomes, it was possible that its diffuse distribution on the sucrose gradients was an artifact during to dissociation during centrifugation. To get a better insight, therefore, we checked the interaction of recombinant aSdo1 with the two purified *S.solfataricus* ribosomal subunits separately. The results clearly show that the factor binds specifically to the 50S subunit, while being incapable of interacting with the purified 30S particle.(Fig.5).

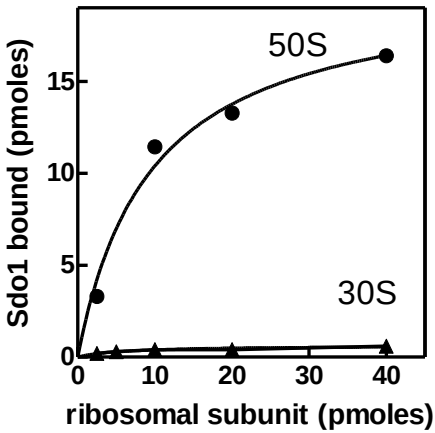

Fig. 5 Interaction of recombinant **aSdo1** with the two purified *S.solfataricus* ribosomal subunits. aSdo1 does not bind 30S subunit while 50S subunit is bound by aSdo1 until about 15 pmol.

3.3 aIF6 gets off translating ribosomes in *S. solfataricus*

Next, we turned to investigate how aIF6 is released from the ribosomes.Since the factor is bound to 50S ribosomes not engaged in translation (as shown by the fact that it is not found on 70S ribosomes), it was reasonable to suppose that it was released once translation had begun. To check this, the *S. solfataricus* whole cell lysate (S30 extract)was incubated at 70°C for 30 minutes under conditions favouring translation. After incubation, the lysate was fractionated on sucrose density gradient and the position of aIF6 was determined by western blotting.As shown in Fig. 6A, in the lysate programmed for translation about half of the total aIF6 amount was found off the ribosomes. Interestingly, the free protein consisted of multiband forms. The reason for this is as yet unknown, but one possibility is that the multiple bands represent unidentified post-translational modifications.

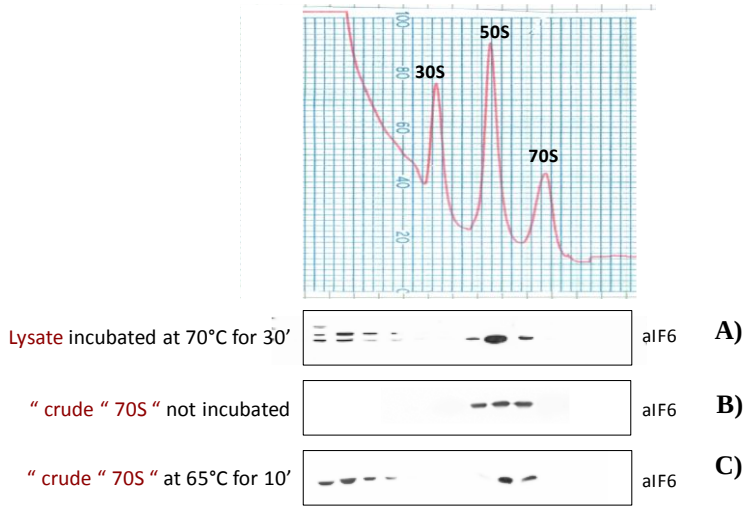

Fig.6 A, Lysate programmed for translation at 70°C for 30 min induces release of aIF6 from the 50S subunits. Furthermore, the free fraction of aIF6 consisted of multiband forms possibly indicative of uncharacterized post-translational modifications. **B**, crude 70S; **C**, in crude 70S, aIF6 is released only after incubation at 65°C for 10'.

More detailed information about the mode of eIF6 release was obtained by *in vitro* experiments in which *S.solfataricus* whole ribosomes, either crude or high-salt purified, were incubated in a variety of conditions. Firstly, we observed that, when crude ribosomes are incubated at 65° C for 10 minutes, a substantial fraction of bound aIF6 is released (Fig. 6B and C). This demonstrates that ongoing translation is not required for aIF6 detachment, but that the ribosomes themselves are competent for it. However, when high-salt washed ribosomes were incubated under the same conditions as before, aIF6 was not released (fig 7B). This suggests that high-salt washing removes some factor essential for aIF6 detachment. That this is indeed the case, as shown by the

fact that adding back the HSW to the purified ribosomes, and incubating the mixture as previously, aIF6 release was again observed (Fig. 7C). Importantly, the aIF6 detachment reaction requires the hydrolysis of GTP. Indeed, crude 70S ribosomes incubated at 65°C for 10 minutes with the addition of the 3 mM GMP-PNP (a non-hydrolyzable GTP analog) blocked the release of aIF6 (Fig.8B). This suggested that, as it happens in the eukaryotes, some GTPase was implicated in removing aIF6 from the ribosome.

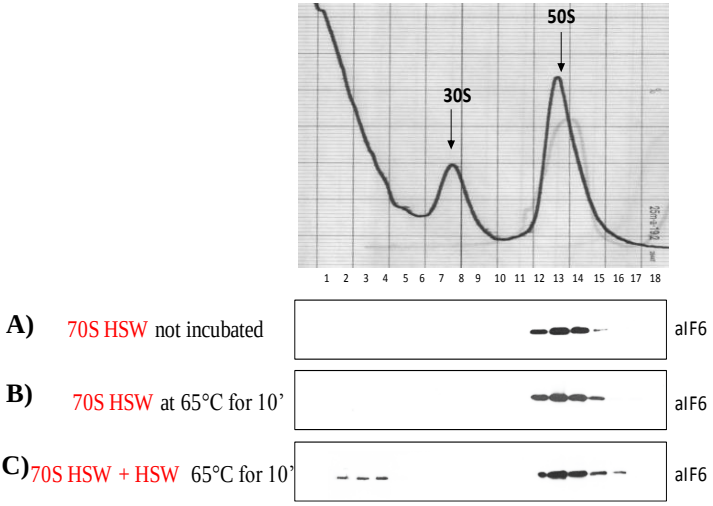

Fig. 7. Incubation of high salt washed ribosomes (70SHSW) in absence of recombinant aSdo1. **A**. 70S HSW non-incubated. **B**. 70S HSW incubated at 65°C for 10'. **C**. 70S HSW with HSW at 65°C for 10'.

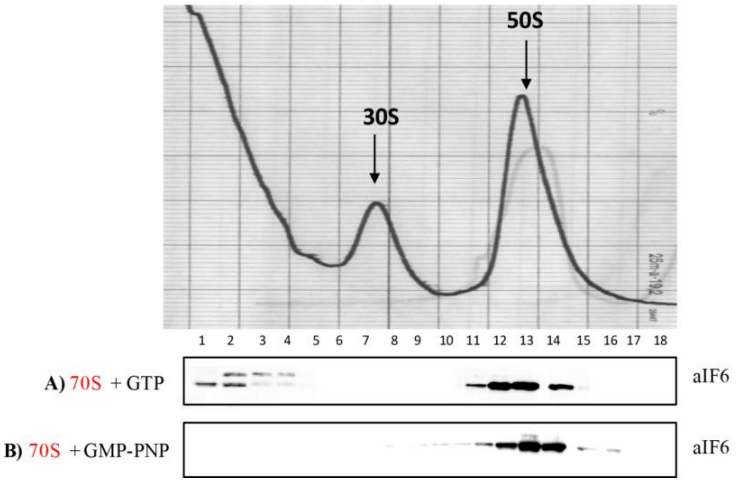

Fig. 8. The GMP-PNP inhibits aIF6 release from crude ribosomes. **A)** 70S + GTP at 65°C for 10'. **B)** 70S + GMP-PNP at 65°C for 10'.

3.4 Sdo1 by itself does not detach aIF6.

In eukaryotes, eIF6 release depends upon the activity of a specialized GTPase, EFL1, which acts together with SBDS/Sdo1. Archaeal genomes, however, do not contain any homologue of EFL1, and in general contain very few putative GTPases. Two questions therefore arose: which was the GTPase required for aIF6 detachment in the Archaea? Was aSdo1 essential for aIF6 release, and did the eukaryotic model for such release apply to the Archaea?

The involvement of aSdo1in the removal of aIF6 was suggested, besides the homology with the corresponding eukarytic factor, also by the fact that HSW ribosomes, which lack aSdo1, are unable to release aIF6. To determine whether aSdo1 could catalyze aIF6 removal, we incubated HSW ribosomes under different conditions, with and without recombinant Sdo1, and checked aIF6 release using both centrifugation on sucrose gradients and size-exclusion chromatography.

The results, shown in Fig. 9, inequivocally show that aSdo1, by itself, does not catalyze aIF6 detachment.

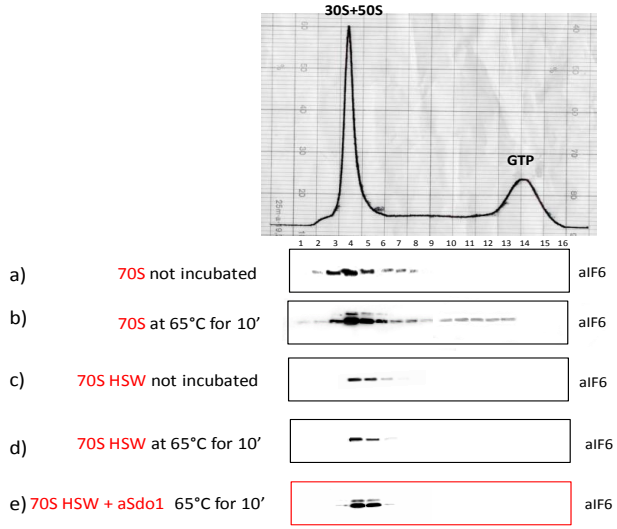

Fig. 9. Analysis of aIF6 binding to the 50S ribosomal subunits by size-exclusion chromatography. **A) 70S** not incubated. **B) 70S** incubated at 65°C for 10'. **C) 70SHSW** not incubated. **D) 70SHSW** at 65°C for 10'. **E) 70SHSW** with aSdo1 at 65°C for 10'.

This result, of course, was not unexpected, since the required GTPase was probably lacking. In the search for this, we preliminarly analyzed the GTPase activities in the various preparations, namely crude and HSW ribosomes, and HSW supernatant. To this end, we performed non-radioactive GTPase assays, where hydrolysis of the nucleotide is assayed by a colorimetric test (see Methods).

As shown in Fig. 10, the crude ribosome have a high GTPase activity, which is, however, lost upon high salt washing. HSW also has, as expected, the capability of hydrolyzing GTP. Interestingly, aSdo1by itself shows no capacity to hydrolyze GTP(Fig. 12), but it is apparently

able to act as a GTPase to some extent when incubated with HSW ribosomes(Fig. 11). We did not investigate further this point, however, since this ribosome-dependent GTPase activity of Sdo1 is apparently not sufficient to cause aIF6 release.

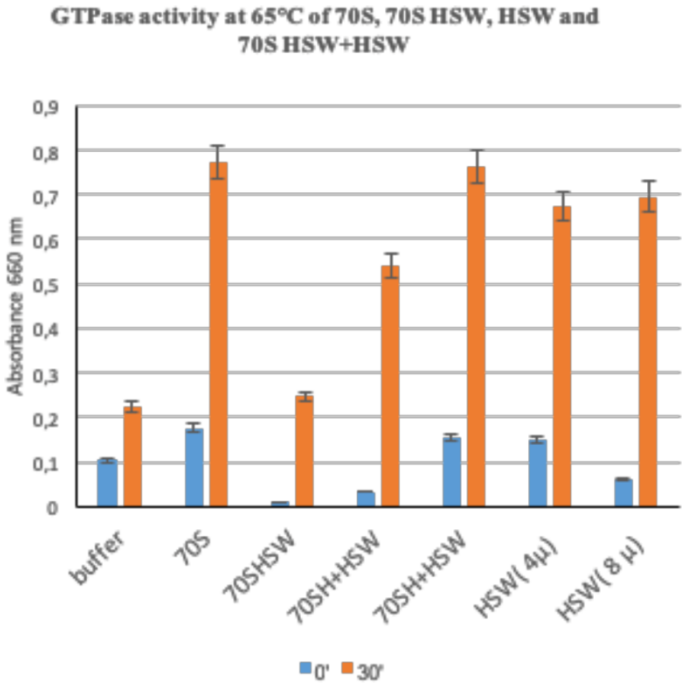

Fig. 10. GTPase assay at 65°C of *S. solfatarucus* fractions. Comparison between the GTPase activity at 65°C of the 70S, 70SHSW ribosomes and HSW.

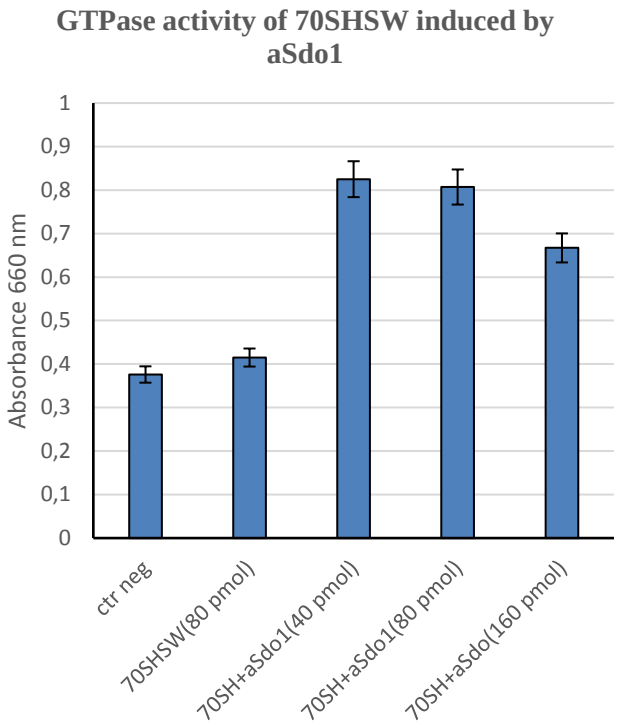

Fig.11. GTPase assay of 70SHSW with aSdo1 protein. GTPase activity at 65°C of 80 pmol 70SHSW plus 40, 80 and 160 pmol aSdo1 respectively.

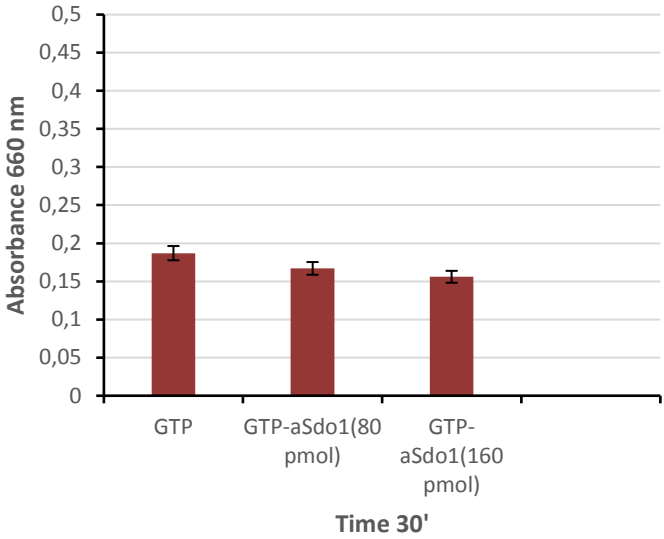

Fig. 12. GTPase activity at 65°C of a Sdo1 protein. Both 80 and 160 pmol aSdo1 have not a GTPase activity.

3.5 Functional characterization of aEF2 protein

As said before, Archaea do not possess homologues of the specialized GTPase EFL1, involved in eIF6 release in eukaryotes. However, EFL1 is a close homologue of elongation factor 2 (EF2), which raised the possibility that, in Archaea, EF2 itself could be the G protein implicated in aIF6 detachment.

To verify this surmise, we decided to clone the gene encoding for the aEF2 protein in our archaeal model organism *Sulfolobus solfataricus*. On the basis of the genome sequence of *S. solfataricus*, a couple of PCR primers were designed to amplify the *aef2* gene by PCR. The product of the reaction was inserted into the expression vector pETM11+, by following the expression/purification of this C-terminus His_6-tagged protein as previously described for Sdo1. Gene expression was induced by IPTG; after 3 h of induction, the amount of expressed protein was sufficient for subsequent purification. The aEF2 protein was purified from *E. coli* extracts under native conditions by affinity chromatography on a Ni^{2+}agarose column that binds selectively the His-tagged protein. (Fig. 13).

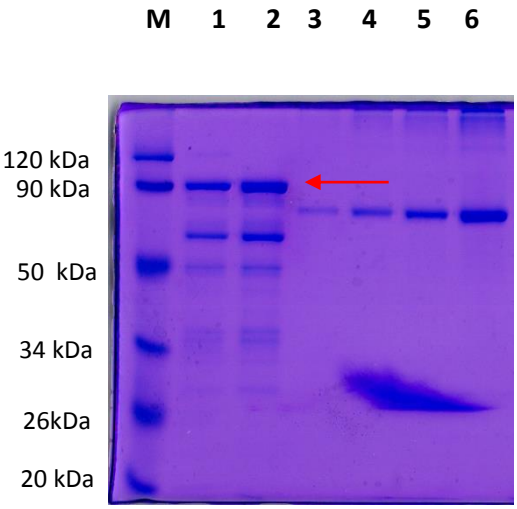

Fig. 13. Expression and purification of aEF2. *aef2* gene was cloned in the p ETM11+ expression vector and the construct was expressed in *E. coli* BL21(DE3) strain. At the end of expression, the recombinant protein was purified by Ni^{2+}agarose columns. The product of purification was loaded on 12,5% SDS-PAGE and stained with Coomassie **M**, molecular weight marker; **1**, with 25 pmol; **2**, with 50 pmol;**3-6,** known amounts of BSA: **3**,0,5 µg;**4**,1 µg;**5**, 2 µg; **6**, 4 µg. The arrow indicates the recombinant protein aEF2.

3.6 aEF2 catalyzes aIF6 release

We checked whether recombinant aEF2 was able to assist aSdo1 in the removal of aIF6 from ribosomes. This analysis was performed by sucrose gradient experiments.The results, illustrated in Fig.14A, revealed that, in the presence of GTP, aEF2 was able to provoke the detachment of aIF6 by itself, in the absence of aSdo1. Similar results were obtained by size-exclusion chromatography (Fig. 15A). Moreover, this analysis was also performed in presence of aEF2 and aSdo1 together. Unexpectedly, addition of recombinant aSdo1 not only did not improve the release of aIF6, but actually inhibited it (Fig.15B).

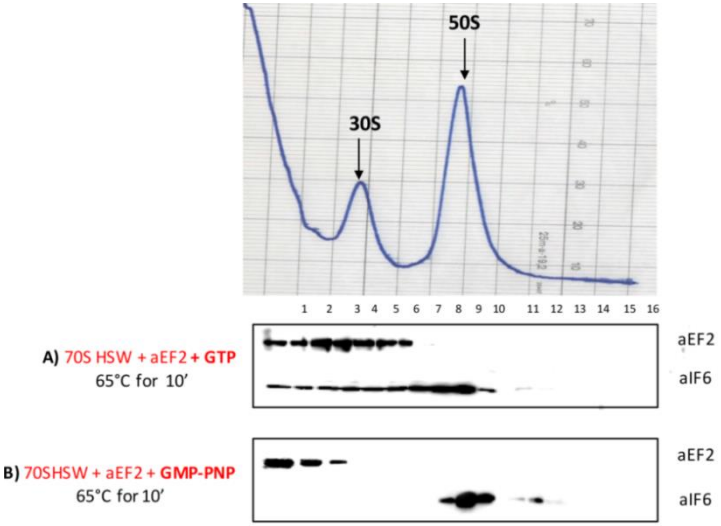

Fig. 14. aEF2 induces *per se* the release of aIF6 from the 50S subunits. **A)** HSW ribosomes incubated with aEF2 and GTP; **B)** HSW ribosomes incubated with aEF2 and GMP-PNP.

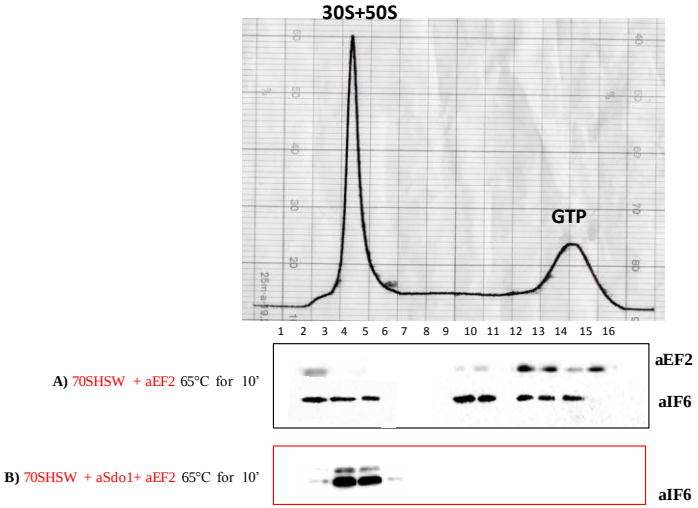

Fig. 15. Analysis of **aEF2** as a GTPase involved in aIF6 release.**A)** Incubated HSW ribosomes + aEF2 + GTP. **B)** Incubated HSW ribosomes+aEF2 + aSdo1+ GTP.

To get a better insight into this unexpected result, we repeated the experiments by incubating the ribosomes with Sdo1 and aEF2, but adding the proteins in a different order. In one sample aIF2 was added first, and Sdo1 after 10 min, while in another sample the reverse ws done. As shown in Fig.16, the addition of aSdo1 before aEF2 effectively inhibited aIF6 release. These results suggest that Sdo1 and aEF2 share, at least in part, a common binding site on the ribosomes, and that the presence of aSdo1 prevents the binding of aEF2 thus preventing release of aIF6.

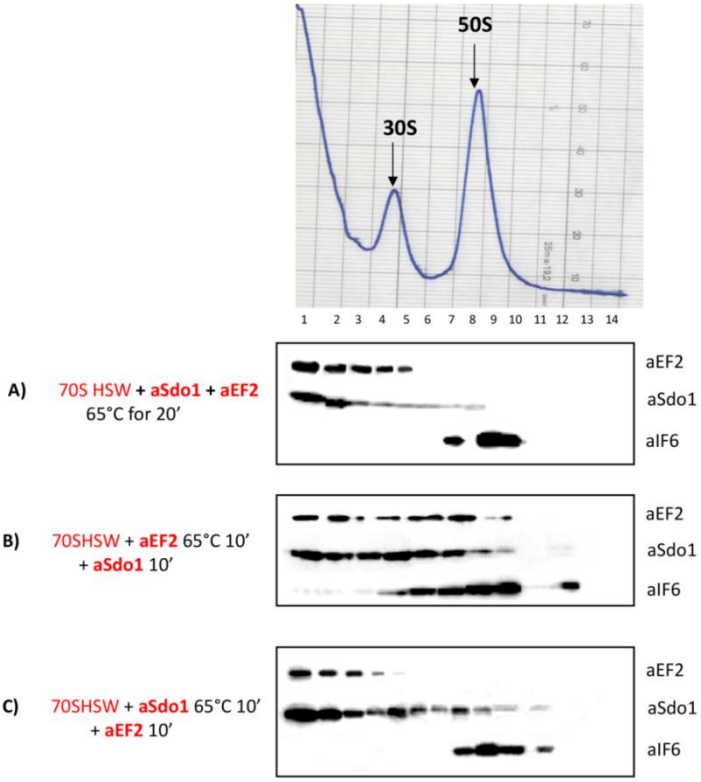

Fig. 16. **aSdo1** inhibits the release of aIF6 induced by **aEF2. A)** HSW ribosomesincubatedwith aSdo1 and aEF2 at 65°C for 20'. **B)** HSW ribosomes incubated before with aEF2 (for 10') and then with aSdo1 for another 10'. **C)** HSW ribosomes incubated before with aSdo1(for 10') and then with aEF2 for another 10'.

4.DISCUSSION

4.1 The release of aIF6 from ribosome

In this work, the mechanism of release of the translation factor aIF6 from the large ribosomal subunit has been experimentally studied for the first time. Although a final mechanism has not been defined and will require further work, the results obtained have unveiled interesting homologies and differences with the corresponding eukaryotic process.

Firstly, we could conclude that aIF6 release from archaeal ribosomes, similar to eukaryotes, is a GTPase-dependent event. The involved GTPase is the elongation factor 2 (aEF2), which by itself is necessary and sufficient to observe aIF6 detachment from the ribosomes, even in the absence of ongoing translation. Since Archaea do not possess a homologue of the GTPase EFL1, involved in the eIF6 release in eukayotes, a similar role of aEF2 in the process had already been suggested on the basis of the fact that EFL1 is a close homologous of eEF2.

We found that, indeed, aEF2 was able to cause the release of aIF6 from the large ribosomal subunit. The reaction depended on the GTPase activity of the factor, showing that also in Archaea the detachment of aIF6 from the ribosomes is dependent by an active, energy consuming, process.
However, the archaeal aIF6 release system appears not to require aSdo1, the SBDS homologue. Instead, aSdo1 seems to have an inhibitory effect on aIF6 detachment, probably because it binds to the ribosomes on a partially competitive aEF2 site. Indeed, also eukaryotic SBDS was shown to share in part the same binding site with the GTPase EFL1; however, in the eukaryotic system, the

arrival of EFL1 causes a conformational change of SBDSthat is in turn required for the ejection of eIF6. This does not seem to be the case in Archaea, even if further experimental work will be necessary to really understand the mechanism for aIF6 release, as well as to understand the role played by aSdo1 in translation.

5.REFERENCES

Bailey K.A., Chow C.S. and Reeve J.N. (1999). Histone stoichiometry and DNA circularization in Archaeal nucleosomes. *Nucleic acids Res*. **27**(2): 532-536.

BarnesS.M., Fundyga R.E., Jefferies M.W. and Pace N.R. (1994). Remarkable Archaeal diversity detected in a Yellowstone National Park hot spring environment. *P.N.A.S. USA* **91**:1609-1613.

Barnes S.M., Delwiche C.F., Palmer J.R. and Pace N.R. (1996). Perspective on Archaeal diversity, thermophily and monophylyl from environmental rRNA sequences. *P.N.A.S. USA* **93**: 9188-9193.

Belfort M. and Weiner A. (1997). *Another* bridge between kingdoms: tRNA splicing in Archaea and eukaryotes. *Cell* **89**: 1003-1006.

Bell S.D. and Jakson S.P. (1998). Transcription and translation in Archaea: a mosaic of eucaryal and bacterial features. *Trends in Microbiology* **6**: 222-228.

Benelli D.,Maone E. and Londei P. (2003). Two different mechanisms for ribosome/mRNA interaction in Archaeal translation initiation. *Mol Microbiol*. **50**(2): 635-43.

Benelli D., Marzi S., Mancone C., Alonzi T., la Teana A. and Londei P. (2009).Function and ribosomal localization of aIF6,a translational regulator shared by Archaea and eukarya. *Nucleic Acids Res.* **37**(1):256-67.

Berchtold H., Reshetnikova L., Reiser C.O., Schirmer N.K., Sprinzl M. and Hilgenfeld R. (1993). Crystal structure of active elongation factor Tu reveals major domain rearrangements.*Nature.* **365**(6442):126-32.

Biffo S., Sanvito F., Costa S., Preve L., Pignatelli R., Spinardi L. and Marchisio P.C.(1997). Isolation of a novel beta4 integrin-binding protein (p27(BBP)) highly expressed in epithelial cells. *JBiol Chem.* **272**(48):30314-21.

Czworkowski J., Wang J., Steitz T.A. and Moore P.B. (1994). The crystal structure of elongation factor G complexed with GDP, at 2.7 A resolution. *EMBO J.* **13**(16):3661-8.

Dennis P.P. (1997). Ancient ciphers: translation in Archaea. *Cell* **89**: 1007-1010.

Finch A.J., Hilcenko C., Basse N., Drynan L.F., Goyenechea B., Menne T.F., González Fernández A., Simpson P., D'Santos C.S., Arends M.J., Donadieu J., Bellanné-Chantelot C., Costanzo M., Boone C., McKenzie A.N., Freund S.M. and Warren A.J. (2011). Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes Dev.***25**(9):917-29.

Grabowski B. and Kelman Z. (2003). Archaeal DNA replication: proteins in a bacterial context. *Annu. Rev. Microbiol*. **57**: 487-516.

Grigoriev A. (1998). Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res*. **26**: 2286-2290.

Groft C.M., Beckmann R., Sali A. and Burley S.K. (2000). Crystal structure of ribosome anti-association factor IF6. *Nat Struct Biol.***7**(12):1156-64.

Grundy F.J. and Henkin T.M. (2006). From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. *Crit Rev BiochemMol Biol*.**41**(6): 329-38.

Gualerzi C.O. and Pon C.L. (1990). Initiation of mRNA translation in prokaryotes. *Biochemistry* **29**: 5881-5889.

Huber H., Hohn M.J., Rachel R., Fuchs T., Wimmer V.C. and Stetter K.O. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature***417**: 63-67.

Kandler O. and Köenig H. (1993). Cell envelopes of Archaea: Structure and chemistry. *The Biochemistry of Archaea(Archaeabacteria) Elsevier Science Publishers B.V.*

Kates M. (1993). Membrane lipids of Archaea in: *The Biochemistry of Archaea(Archaeabacteria) Elsevier Science Publishers B.V.*

Keeling P.J. and Doolittle W.F. (1995). Archaea: Narrowing the gap between prokaryotes and eukaryotes. *P.N.A.S. USA* **92**: 5761-5764.

Keeling P.J., Charlebois R.L. and Doolittle W.F. (1994). Archaebacterial genomes: Eubacterial form and eukaryotic content. *Current opinions in Genetics and Development* **4**: 816-822.

Kyrpides N.C. and Woese C.R. (1998). Universally conserved translation initiation factors. *P.N.A.S. USA* **95**: 224-228.

Klinge S., Voigts-Hoffmann F., Leibundgut M., Arpagaus S. and Ban N. (2011). Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6.*Science.* **334**(6058):941-8.

Kozak M. (1983). Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbiol. Rev*. **47**: 1-45.

Langer D., Hain J., Thuriaux P. and Zillig W. (1995). Transcription in Archaea: Similarity to that in Eucarya. *P.N.A.S. USA* **92**:5768-5772.

Lebreton A.,Saveanu C., Decourty L., Rain J.C., Jacquier A. and Fromont-Racine M. (2006). A functional network involved in the recycling of nucleocytoplasmic pre-60S factors. *J Cell Biol.* **173**(3):349-60.

Lempiäinen H. and Shore D. (2009). Growth control and ribosome biogenesis. *CurrOpin Cell Biol.* **21**(6):855-63.

Lykke- Andersen J., Aagaard C., Semionenkov M., Garrett R.A. (1997). Archaeal introns: Splicing, intercellular mobility and evolution. *Trends in Biochemical Sciences* **22**:326-331.

Lundgren M., Andersson A., Chen L., Nilsson P. and Bernander R. (2004). Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *P.N.A.S. USA* **101**: 7046-7051.

Maisnier-Patin S., Malandrin L.,Birkeland N.K. and Bernander R. (2002). Chromosome replication patterns in the hyperthermophiliceury Archaea *Archaeoglobus fulgidus* and *Methanocaldococcus (Methanococcus) jannaschii*. *Mol. Microbiol.* **45**:1443-1450.

Martin W. and Koonin E.V. (2006). A positive definition of prokaryotes.*Nature.* **440**: 41-45.

Matsunaga F., Forterre P., Ishino Y. and Myllykallio H. (2001). *In vivo* interactions of archaeal Cdc6/Orc1 and minichromosome maintenance protein with the replication origin. *P.N.A.S. USA* **98**:11152-11157.

Myllykallio H., Lopez P., Lopez-Garcia P., Heilig R., Saurin W., Zivanovic Y., Philippe H. and Forterre P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**:2212-2215.

Ouzounis C. and Kyprides N. (1996). The emergence of major cellular processes in evolution. FEBS *Letters* **390**: 119-123.

Pereira S.L., Grayling R.A., Lurz R. and Reeve J.N. (1997). Archaeal nucleosomes.*P.N.A.S.***94**(23):12633-7.

Pestova T.V. and Hellen C.U. (2000). The structure and function of initiation factors in eukaryotic protein synthesis. *Cell Mol Life Sci*. **57**(4):651-74.

Pestova T.V. and Kolupaeva V.G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev*. **15**: 2906-2922.

Robinson N.P., Dionne I., Lundgren M., Marsh V.L., Bernander R. and Bell S.D. (2004). Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus. Cell* **116**: 25-38.

Rodionov D.A., Vitreschak A.G., Mironov A.A. and Gelfand M.S. (2003). Regulation of lysine biosynthesis

and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res.* **31**(23): 6748-57.

Sachs A.B., Sarnow P. and Hentze M.W. (1997). Starting at the beginning, middle and end: translation initiation in eukaryotes. *Cell* **89**: 831-838.

Salzberg S.L., Salzberg A.J., Kerlavage A.R. and Tomb J.F. (1998). Skewed oligomers and origins of replication. *Gene* **217**:57-67.

Santangelo T.J. and Reeve J.N. (2006). Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. *J MolBiol* **355**(2):196-210.

Sanvito F., Piatti S., Villa A., Bossi M., Lucchini G., Marchisio P.C. and Biffo S. (1999). The beta4 integrin interactor p27(BBP/eIF6) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly. *J Cell Biol.* **144**(5):823-37.

Schwartz R.M. and Dayhoff M.O. (1978). Origins of prokaryotes, eukaryotes, mitochondria and chloroplasts. *Science* **199**: 395-403.

She Q., Singh R.K., Confalonieri F., Zivanovic Y.,Allard G., Awayez M.J., Chan-Weiher C.C., Clausen I.G., Curtis B.A., De Moors A., Erauso G., Fletcher C., Gordon P.M., Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X., Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N., Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T., Garrett R.A., Ragan M.A., Sensen C.W. and Van der Oost J. (2001). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *P.N.A.S. USA* **98**(14): 7835-7840.

Torarinsson E., Klenk H.P. and Garrett R.A. (2005). Divergente transcriptional and translational signals in Archaea. *Environ Microbiol*. **7**(1):47-54.

Woese C.R. and Fox G.E. (1997). Phylogenetic structure of the prokaryotic domain: The primary Kingdoms. *P.N.A.S. USA* **74**: 5088-5090.

Woese C.R., Kandler O. and Wheelis M. (1990). Towards a natural system of organisms: Proposal from the domains Archaea, Bacteria and Eucarya. *P.N.A.S. USA* **87**:4576-4579.

Valenzuela D.M., Chaudhuri A. and Maitra U. (1982). Eukaryotic ribosomal subunit anti-association activity of calf liver is contained in a single polypeptide chain protein of Mr = 25,500 (eukaryotic initiation factor 6). *J BiolChem*. **257**(13):7712-9.

Volta V., Ceci M., Emery B., Bachi A., Petfalski E., Tollervey D., Linder P., Marchisio P.C., Piatti S. and Biffo S. (2005). Sen34p depletion blocks tRNA splicing *in vivo* and delays rRNA processing.*BiochemBiophys Res Commun*. **337**(1):89-94.

Chapter 2

Optimization of an *in vitro* transcription/translation system based on *Sulfolobus solfataricus* cell lysate.

1.INTRODUCTION

1.1 *In vitro* transcription/translation system

At the beginning of the 1950s, several groups independently demonstrated that protein synthesis does not require the integrity of the cell and can continue after cell disruption. Thus, disrupted cells or their isolated fractions were reported to be capable of synthesizing proteins (Gale and Folkes, 1954). In the meantime, ribonucleoproteinparticles were observed and identified in cells and then isolatedfrom cells and studied with respect to their physicochemical properties (Peterman and Hamilton, 1957). The protein- synthesizing ability of these particles was experimentally proved (McQuillen *et al*,1959). The word "ribosome" wasproposed to designate the protein-synthesizing ribonucleoprotein particles.

In the second half of that decade, researches conducted cell-free protein synthesis based on mitochondria-free cytoplasmicextracts of animal cells (Keller and Zamecnik, 1956;Littlefield and Keller, 1957). The dependence of the system on energy supply in theform of ATP and GTP was shown.

Other researches made a real ribosomalsystem of protein synthesis (translation) based on human and rabbit reticulocytes (Bank and Marks, 1966), ascites cells (Keller and Littlefield, 1957) and wheat germ (Marcus and Feeley, 1966). Ribosomes in all ofthose systems were programmed with endogenous mRNA; they were simply readingthe messages to which they had already been attached at the time of cell disruption.

Nevertheless, the significance of these systems was great, since theyopened the door for studies of molecular mechanisms of protein biosynthesis, including activation of amino acids, involvement of tRNA, GTP requirement, ribosome function and participation of soluble translation

factors (Zamecnik, 1969). Since the preparation processes did not remove messenger RNA (mRNA), the cellextracts that comprised early cell-free systems synthesized mostly native proteins from endogenous mRNA. These extracts were essentially supernatants obtained by centrifugation of lysates from disrupted cells e.g., at 30,000*g* (the so-called S30 extract) (Schweet *et al.*, 1958).

To direct the protein synthesis machinery to translate exogenous mRNA, the *E. coli* S30 extract were incubated in the presence of 20 amino acids and ATP regeneration components (Nirenberg and Matthaei, 1961). Such "incubated S30" improved the amount of the protein synthesized from exogenous mRNA, possibly by freeing ribosomes from attaching to endogenous mRNA (ribosome run-off) and allowing endogenous mRNA to be degraded by ribonucleases in the extract.

Later, the coupled transcription-translation systems were introduced in cell-free protein synthesis using DNA as the template instead of mRNA (Chen and Zubay, 1983; Zubay, 1973). For istance, the *E. coli* RNA polymerase was used to transcribe DNA into mRNA for protein synthesis in the S30 extracts, mainly for the purpose of studying gene regulation *in vitro*. (Josephsen and Gaastra, 1985). To obtain higher amounts of mRNA, the researchers later used stronger phage promoters and the more efficient phage RNA polymerases (T7 or SP6) for the coupled transcription-translation from the DNA template (Craig *et al*., 1992; Krieg and Melton, 1987). In addition to using the more stable DNA template and avoiding a separate *in vitro* transcription step, coupled transcription and translation in the cell-free system often result in a higher protein synthesis yield, probably due to newly synthesized mRNA is immediately translated into

protein, which may minimize the adverse effects of mRNA degradation or inhibitory structures.

Cell-free protein synthesis is one of widely used methods in molecular biology. Production of proteins using cell-free protein synthesis usually takes a few hours, in contrast to production of proteins in cells, which typically takes days if not weeks. In fact, even first-time users can often obtain newly synthesized proteins on the same day he or she begins to use a commercial system.

Cell-free protein synthesis (CFPS) systems derived from crude cell extracts have been used for decades as a research tool in fundamental and applied biology, as illustrated in Fig.1.

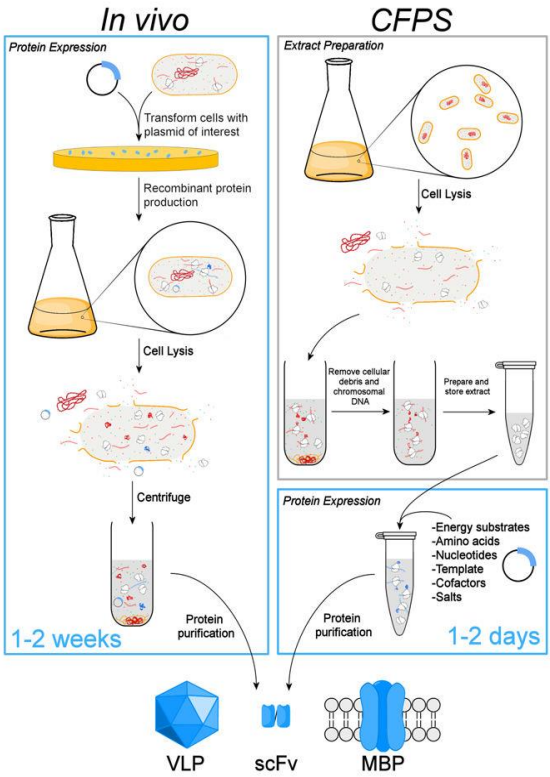

Fig.1.Cartoon comparison of *in vivo* recombinant DNA protein expression with cell-free protein synthesis (CFPS). CFPS systems provide a more rapid process/product development timeline.
Example proteins shown include a virus-like particle (VLP), single-chain antibody variable fragment(scFv) and a membrane bound protein (MBP).

The use of cell-free protein synthesis has made the most impact on functional and structural genomics. For the first time, researchers have been able to express and purify a large number of proteins in a short period of time for subsequent high throughput functional and structural analyses. For instance, a number of laboratories have used low-cost *E. coli* extract and wheat germ cell-free

systems in high-throughput automated format to produce stable isotope-labeled proteins for nuclear magnetic resonance analyses (Aoki *et al*., 2009; Vinarov and Markley, 2005). The advantage is that only newly synthesized proteins are labeled during cell-free protein synthesis and can be analyzed without extensive purification. Moreover, CFPS has shown remarkable utility as a protein synthesis technology (Swartz, 2006), including the production of pharmaceutical proteins (Yang *et al.*, 2005), and high-throughput production of protein libraries for protein evolution and structural genomics (Takai *et al*.,2010*).* In particular, cell-free systems havedistinct advantages over *in vivo* methods for recombinant protein production (Zawada *et al.*, 2011).
Without the need to support ancillary processesrequired for cell viability and growth, CFPS allows optimization of the cell extract towards the exclusive production of asingle protein product. The absence of a cell wall enables an open and versatile environment for active monitoring, rapid sampling, and direct manipulation of the protein synthesis process.
In Archaea, cell-free protein synthesizing systems have been in use for a couple of decades, and proven to be of great utility for understanding the basic features of translation, as weel as for the synthesis of thermostable proteins. Most CFPS from Archaea, however, worked optimally with pre-synthesized mRNAs. Here, I report the development of a coupled *in vitro* transcription/translation system for cell-free protein synthesis from the thermophilic archaeon *S.solfataricus*. This system permits the efficient *in vitro* synthesis of proteins at high temperature and it is based on the use of an unfractionated cell lysate (S30) adapted to perform coupled transcription and translation. First at all, it

represents a powerful tool to expand our understanding of the molecular mechanisms governing coupled transcription-translation in Archaea. Moreover, the expression of recombinant proteins in thermophilic conditions similar to the native ones could facilitate the identification of associated factors. Furthermore, although mesophilic hosts such as *Escherichia coli* have been used to produce thermostable proteins for biochemical and crystallographic characterization (Ward *et al.*, 2000; Watanabe *et al*., 2010), many hyperthermophilic proteins correctly fold only under physiological conditions of high temperature or in the presence of their native post-translational modifications. (Andreotti *et al*., 1995).

2.MATERIALS AND METHODS

2.1 Gene constructs and *in vitro* transcription

We used the plasmid pBluescript-SK(+) as a starting point for our subsequent constructs. Two synthetic DNA oligomers of 48 nucleotides were designed on the conserved sequence of *S. solfataricus* 16S/23S rRNA operon promoter: Promoter rRNA SSO Forward 5'-CGAAGTTAGATTTATATGGGATTTCAGAACAAT ATGTATAATGGGTAC-3' and Promoter rRNA SSO Reverse 5'-CCATTATACATATTGTTCTGAAATCCCATATAAA TCTAACTTCGGTAC-3'. After annealing of the two oligomers, one pmol of the purified double strand fragment was incubated with 0.25 pmol of *Kpn I* digested pBS-SK(+) plasmid in the presence of 10 units of T4 DNA ligase (New England BioLabs) in 25 µl of 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 µg/ml bovine serum albumin for 20 h at 16 °C. One tenth of this reaction mixture was then used directly for transformation of *E.coli* Top 10 competent cells. The clone harbouring the construct with the insert in the correct orientation was selected after DNA sequencing and termed pBS-$rRNA_p$. Successively, a fragment of 393 bp containing the gene termed ORF 104 with its Shine-Dalgarno (SD) motif was amplified from the construct pBS800 (Condò *et al.*, 1999) by PCR using the following primers: Prom-104 *Xho I* 5'-TTTTTTTATCTCGAGCCGGAATAGTTGAATTAAC AATGAAGC-3' and Prom-104 *Pst I* 5'-CATGGTATGCTGCAGTCATTGCTTCACCTCTTTA ATAAACTCC-3'. The fragment was inserted into the *Xho I-Pst I* digested plasmid pBS-$rRNA_p$, yielding the construct termed pBS-$rRNA_p$-104. To generate the construct termed pBS-$rRNA_p$-OGT, the fragment *Xho I-*

Pst I was excised from the previously plasmid and inserted a DNA fragment of 533 bp amplified from the construct pQE-*ogt* (Perugino *et al*.,2012) by PCR with the following primers: Forward rRNA/ SsOGT *Xho I* 5'-TTTTTCTCGAGTGAGGTGAAATGTAAATGAGAG GATCTCACCATCACC-3' and Reverse rRNA/ SsOGT *Pst I* 5'-TTTTTTCTGCAGTCATTCTGGTATTTTGACTCCC-3'.

2.2 *In vitro* labeling of transcriptional activity of *Sulfolobus solfataricus* lysate

The transcriptional activity of the *S. solfataricus* cell-free extract was tested by ^{32}P-UTP incorporation in two different reaction conditions using an aliquot of the lysate corresponding to 100 μg of total proteins. In the first reaction, the cell-free extract was incubated in a reaction volume of 50 μl, in the presence of 50 mM Tris/HCl (pH 8.0), 25 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol, 2 mM ATP, 1 mM GTP, 1 mM CTP, 0.6 mM UTP and 100 μM [α-^{32}P] UTP (4 Ci/mmol) in a reaction volume of 50 μl. The reaction was carried out at 60° C for 30 min. The second reaction was performed on the basis on the *in vitro* translation experiments carried out previously (Benelli and Londei, 2007): *S. solfataricus* cell-free extract was incubated in a reaction volume of 50 μl, in the presence of 10 mM KCl, 20 mM Tris/HCl (pH 6.8), 20 mM $Mg(OAc)_2$, 2 mM ATP, 1 mM CTP, 1 mM GTP, 0.5 mM UTP and 100 μM [α-^{32}P] UTP (4 Ci/mmol). The reaction, in this case, was carried out at 70° C for 30 min. At the end of both reactions 20 U of DNase I were added

and incubation was extended for 30 min at 37 °C. The products of the reactions were extracted by phenol pH 4.7 and precipitated with 2.5 volumes of 95% ethanol. The pellets were resuspended in an adequate volume of DEPC-treated water and divided in two aliquots. RNase A (20 μg) was added to one of them and both aliquots were incubated at 37 °C for 30 min. The newly synthesized RNA was separated by 8,5% of non-denaturing polyacrylamide gels and detected using both an Istant Imager apparatus (Pakard) and autoradiography film (Kodak XAR-5).

2.3 *In vitro* translation and coupled *in vitro* transcription-translation

The transcription-translation activity was measured in a final volume of 25 μl and contained: 10 mM KCl, 20 mM Tris/HCl (pH 6.8), 20 mM $Mg(OAc)_2$, 1.5 mM ATP, 1.5 mM CTP, 1.5 mM GTP, 1.5 mM UTP, 3,3 μg of bulk *S. solfataricus* tRNA, 5 μl of 20-25 mg/ml *S. solfataricus* S30 extract (preincubated for 10 min at 70 °C) and 0,5 μl of L-[^{35}S]-Methionine (S.A. 1175 Ci mmol^{-1} at 11 mCi ml^{-1}, PerkinElmer). After mixing all components, 4 μg of the desired mRNA or different amounts of plasmid were added and the mixtures were incubated at 70 °C. Whole cell lysates were programmed for *in vitro* translation with transcripts of *S. solfataricus* genes ORF 104 and *Ss*OGT cloned in pBS-SK (+) plasmid downstream of T7 RNA polymerase promoter in the presence of ATP and GTP to the final concentration of 1.8 and 0.9 mM, respectively. The analysis of the translation products was performed

by loading 15 µl of the incubation mixture in 16% polyacrilamide/SDS gels; after the run, the gels were dried and autoradiographed.

2.4 qPCR and RT-PCR *Ss*OGT labeling

At the end of *in vitro* transcription or coupled *in vitro* transcription-translation, total RNA was extracted twice from the reactions by phenol reagent (pH 4.7) and precipitated by adding of 2.5 volumes of 95% ethanol.
The pellets were resuspended in an adequate volume of DEPC-treated water and treated with 2 U of DNase I, RNase-free (ThermoFisher Scientific) in an appropriate buffer at 37 °C for 45 min.
0.5 µg of total RNA was retrotranscribed for relative qRT-PCR analysis (SensiFAST™ cDNA Synthesis Kit, Bioline). qPCR was performed with the Applied Biosystem StepOne Real-Time PCR System (ThermoFisher Scientific) using 1/20 of cDNA and 10 µl of GoTaq® qPCR Master Mix (Promega) in a final volume of 20 µl. Cycling parameters were: 95 °C for 2 min, followed by 40 cycles of denaturation at 95 °C for 3 secs, annealing/extension at 60 °C for 30 sec.
Each mRNA expression level was calculated by 2-ΔΔ Ct method and normalized to endogenous aIF6 mRNA. qRT-PCR reaction were performed using SYBER green incorporation (Promega). All samples were done in triplicate and each condition was repeated three times.
The following primer sequences were used for qRT-PCR: aIF6 Forward 5'-ATAAGCGGTAACGATAACGG-3' and aIF6 Reverse 5'-

AATCCCTTAGATTCTCCTTCAG-3'), for pBS (FORWARD 5'-TGGTAACAGGATTAGCAGAG-3' and Reverse 5'-ACCAAATACTGTCCTTCTAGTG-3'). Also for semi-quantitative RT-PCR, total RNA was extracted from the mix reaction.
2 μg of total RNA were retrotranscribed in a final volume of 25 μl with 200U M-MLV reverse transcriptase in 20 μl of mixture reaction for 1 h at 42 °C according to instructions of the supplier (Promega). The reaction contained 1 μM of the followed reverse primer: 5'-GGTTTCCCGACTGGAAAGCGGGCAG-3'. At the end of the reaction, the final volume of the mixture reaction was adjusted to 50 μl and one-tenth of the RT reaction was PCR amplified with Taq DNA polymerase (Promega) for 30 secs at 95 °C, 30 secs at 60 °C and 45 sec at 74 °C (25 cycles) with a final extension step for 7 min at 74 °C. Reverse primers for PCR amplification were the same used in the RT reaction coupled with the following forward primers: 5'-CGAATTCCTGCAGCCCGGGGGATCC-3'. The products of the reactions were separated by agarose-gel electrophoresis and detected by ethidium-bromide staining.

2.5 *Ss*OGT*in vitro* labeling

In vitro expressed OGT was analyzed incubating 8 μg of pBS-rRNA$_p$-OGT plasmid or 200 ng of recombinant *Ss*OGT with 200 μg of *S. solfataricus* whole cell extract in presence of 2.5 μM SNAP-Vista Green™substrate (hereinafter BG-FL) at 70 °C for 60 min. Reactions were stopped by denaturation and samples were subjected to SDS-PAGE, followed by fluorescence *gel-imaging*

analysis using a VersaDoc 4000™ system (Bio-Rad) by applying as excitation/emission parameters a blue LED/530 bandpass filter. For western blot analysis, proteins were transferred onto PVDF filters (Bio-Rad) using the Trans-Blot1 Turbo™ Blotting System (Bio-Rad). The presence of OGT protein was revealed using polyclonal antibodies raised in rabbit against *S. solfataricus* OGT as primary antibodies; the goat anti-rabbit IgG-HRP (Pierce) as secondary antibody and the Amersham Biosciences ECL Plus kit. Filters were incubated, washed and developed according to manufacturer's instructions. Chemiluminescent bands were revealed using a VersaDoc apparatus (Bio-Rad).

3.RESULTS

3.1Analysis of *in vitro* transcription in the S30 fraction of *S. solfataricus*.

Preliminary experiments wereperfo rmed in order toverify whether the whole cell lysate of *S. solfataricus* prepared according to described protocols (Benelli and Londei., 2007) was competent for *in vitro* transcription. Specifically, we compared the transcriptional activity of our system with that of a previously described *Sulfolobus in vitro* transcription assay (Reiter *et al*., 1990) testing the capacity of the S30 extract to incorporate α-^{32}P-UTP. Salt and temperature conditions of the reactions are summarized in Table 1 and described in detail in Materials and Methods.

	In vitro transcription adopted	*In vitro* transcription under our conditions	Coupled *in vitro* transcription-translation	*In vitro* translation
KCl (mM)	-	10	10	10
Tris/HCl (mM)	50 (pH 8.0)	20 (pH 6.8)	20	20
$Mg(OAc)_2$ (mM)	25	20	20	20
ATP (mM)	2	2	1.5	1.8
CTP (mM)	1	1	1.5	-
GTP (mM)	1	1	1.5	0.9
UTP (mM)	0.6	0.5	1.5	-
[α-^{32}P] UTP(μM)	100	100	-	-
EDTA (mM)	1	-	-	-
DTT (mM)	1	-	-	-
Total tRNA (μg)	-	-	3,3	3,3
S30 (μg)	100-150	100-150	100-150	100-150
T (°C)	60	70	70	70

Table 1: Experimental conditions adopted for reactions with S30 *S.solfataricus*

In both cases, we implemented the reactions with the nucleoside triphosphates at the final concentration of 1 mM each (except ATP to 2 mM) and the S30 fraction was prepared omitting DNase I treatment of lysate. (As shown in Fig. 1), both S30 extracts showed the ability to recruit labeled uridine triphosphate supporting the idea that endogenous RNA polymerase was active. However, the extract prepared according to our protocol had a higher efficiency of uridine triphosphate incorporation. Successively, based on a study characterizing the promoter for the single-copy 16S/23S rRNA gene cluster of the extremely thermophilic archaebacterium *Sulfolobus*(Qureshi *et al*., 1995), we cloned this promoter into the pBS-SK(+) plasmid. The construct contained the region of DNA upstream from the transcription start site of the 16S/23S rDNA gene spanning from -1 to -40 bp(Fig. 2a). The plasmid was incubated with the S30 extract and its transcription was analysed by RT-PCR, using primers annealing to a specific region of the plasmid downstream of the cloned gene, thus excluding amplification of the endogenous target. The results showed an efficient transcription of the plasmid following incubation at 70 °C (as shown in Fig. 2b).

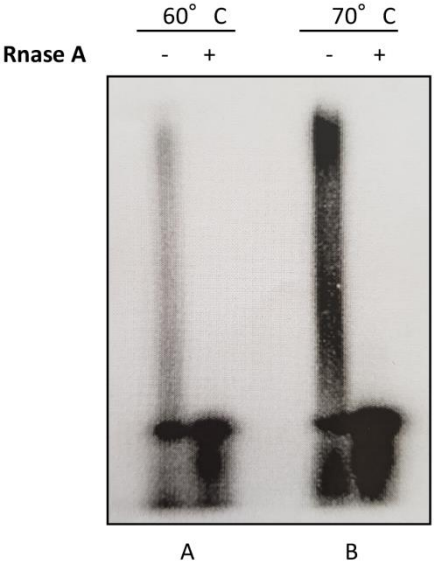

Figure 1: Transcriptional activity of *S. solfataricus* whole cell extracts. *In vitro* transcription reactions were performed using *S. solfataricus* S30 fractions with [α-^{32}P] UTP in different experimental conditions, as described in Material and Methods and Table 1. Reaction A was incubated at 60 °C while reaction B at 70 °C. Total RNA was extracted from the reaction mixes and an aliquot of the samples was treated with Rnase A at 37°C for 30 min. The products of *in vitro* transcription were subjected to non-denaturing polyacrylamide gel electrophoresis and those incorporating [α-^{32}P] UTP visualized by autoradiography.

Starting from this construct, we cloned a previously well-characterized *Sulfolobus* gene encoding a putative ribosomal protein (Condò *et al*., 1999), under the transcriptional control of the 16S/23S rDNA promoter. The structure of this plasmid, termed pBS-rRNA$_p$-104, is shown schematically in Fig. 2c; analysis by qPCR showed that it was also transcribed (Fig. 2d). Finally, the pBS-rRNA$_p$-104 construct was transcribed *in vitro* with T7 RNA polymerase and known amounts of the corresponding purified RNA were used to draw a calibration curve, which was used to quantify the transcription reactions (as illustrated in Fig. 2e).

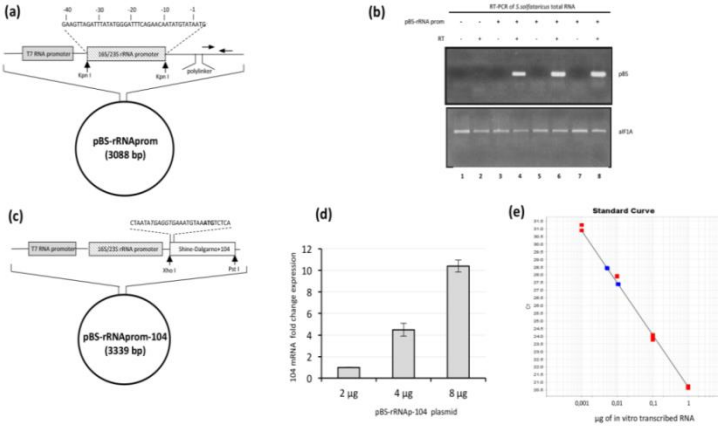

Figure 2: *In vitro* trascription of plasmids containing the 16S/23S rRNA promoter. (a) Schematic representation of pBS-rRNA$_p$ construct. Horizontal arrows indicate the position of primers used for RT-PCR analysis. (b) RT-PCR on total RNA extracted from S30 of *S. solfataricus p*reviously incubated with 4 µg of pBS-rRNA$_p$ plasmid, showing the amplified fragment of 346 bp. (c) Schematic representation of pBS-rRNA$_p$-104 plasmid. The SD motif is evidenced in italic, while the start codon is shown in bold. (d) Relative amount of RNA transcribed by pBS-rRNA$_p$-104 plasmid incubated into *Sso* S30 extract at 70°C for 1h. (e) Absolute quantification of pBS-rRNA$_p$-104 transcript using the standard curve method. The absolute quantities of the standards were obtained measuring the concentration of T7 *in vitro* transcribed pBS-rRNA$_p$-104 RNA. Serial dilutions of the *in vitro* transcript were obtained and their Ct values (red dots) were compared to those unknown (blue dots) extrapolating the amount of copies expressed.

This analysis permitted us to assess the amount of *in vitro* transcribed RNA to an order of magnitude corresponding to ng of RNA for μg of plasmid used, in 25 μl of reaction.The absolute amount of RNA transcribed from the plasmid pBS-rRNA$_p$-104 obtained after its incubation in the in vitro transcription-translation system, was measured by performing RT-qPCR as described above and, then, comparing the Ct values obtained from these samples respect to a standard curve plotted with Ct values obtained serial dilutions of 1 μg of *in vitro* transcribed RNA (pBS-rRNA$_p$-104).Controls correspond to reactions performed on RNA purified from samples in absence of the plasmid and from RT minus cDNA reactions.

3.2 Optimization of *in vitro* translation conditions with respect to NTPs and Mg^{2+} ions

Next, we investigated whether the conditions adopted for *in vitro* transcription with the *S. solfataricus* S30 extract could affect its translational activity. Specifically, we sought to define an optimal concentration of NTPs since it is well known that free nucleotides chelate a proportional number of Mg^{2+} ions, whose presence in a well-defined range of concentration is essential for translation (Nierhaus., 2014). Surprisingly, in a system which contains onlyMg^{2+} and monovalent cations such as K^+ or NH_4^+, the ribosomes will become inactivated at Mg^{2+} concentrations below 10 mM; polyamines, mainly spermidine (2 mM), are necessary in order to rescue the activity and allow the protein synthesis rate even to approach *in vivo* perfection.The optimized systems contain 2 to 5 mM Mg^{2+}. It was demonstrated a genetic link between intracellular Mg^{2+} concentration and

ribosome amounts per cell. (Akanuma *et al*., 2014).In other words, ribosomes represent an important reservoir of the total Mg^{2+} amount, and a reduction of the total ribosome content (in particular, 70S ribosomes and polysomes) by genetic lesions, such as deletions of ribosomal proteins, seems to affect the cellular Mg^{2+} content, affecting probably the free Mg^{2+} concentration. The latter assumed effect,in turn, impairs ribosomal subunit association and thus hampers protein synthesis. So, in order to define the best conditions for *in vitro* translation,we incubated the S30 extract with pre-transcribed 104 mRNA depending from different concentrations of NTPs, and determined its translational efficiency. Indeed, increased levels of NTP in the mix reactions were detrimental for *in vitro* translation (Fig. 3a).

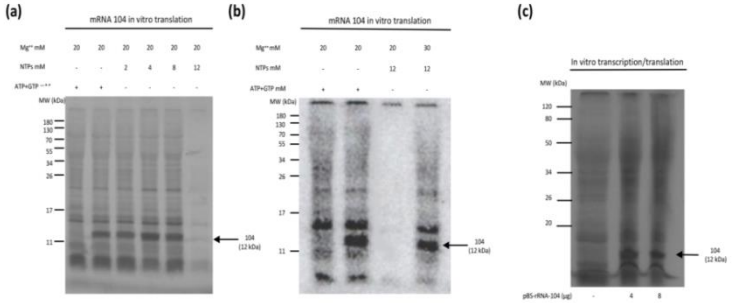

Figure 3: *In vitro* expression of ORF 104. 4 µg of *in vitro* transcribed 104 mRNA were translated at different concentrations of NTPs (a) and Mg^{2+} (b) for 1h in 25 µl of reaction. (c) Different amounts of pBS-rRNA$_p$-104 plasmid were incubated with *S. solfataricus* whole cell extract for 60 min at 70°C in a final volume of 25 µl.

However, this could be in part compensated by increasing the concentration of Mg^{2+} ions as shown in

Fig. 3b. On the other hand, dispensing with added NTPs in the mix reaction completely inhibited the activity of the system, since exogenous ATP and GTP are required as an energy source as illustrated in Fig. 3b, lane 5. Overall, based on the results of Fig. 3a and 3b, we chose to strike a balance between NTP and Mg^{2+} setting them at the final concentration of 6 and 20 mM, respectively.

3.3 Transcription and translation-coupled protein synthesis

We then proceeded to verify whether the previously established experimental conditions allowed coupled transcription and translation. This question was addressed incubating different amount of the pBS-$rRNA_p$-104 plasmid with the lysate at 70 °C for 1h under the conditions summarized in Table 1. The predicted mRNA was endowed with a 5'-UTR containing a SD motif 7 nucleotides upstream from the AUG start codon of ORF 104. As shown in Fig. 3c, the reaction yielded of a main protein band of about 12 kDa, corresponding to the expected size of the ORF 104.

To extend the above results to other *S. solfataricus* genes, we sub-cloned the O^6-DNA-alkyl-guanine-DNA-alkyl-transferase gene (*Ss*OGT) from the pQE-*ogt* construct, previously characterized by Perugino and colleagues (Perugino *et al*., 2012).The product of this gene is a thermostable protein of about 17 kDa, belonging to an evolutionary conserved class of proteins involved in the direct repair of DNA lesions caused by alkylating agents. DNA alkyl-transferases (AGTs or OGTs), catalyze this repair by a one-step irreversible mechanism, involving

the transfer of the alkylic group fromO^6-alkyl-guanine orO^4-alkyl-thymine to a cysteine residue in their own active site. For these reasons, they are classified as "suicide enzymes". On the other hand, the peculiar irreversible reaction of AGTs, led to the setting up of a new *protein-tag*, the so-called SNAP-*tag*™, which represents an alternative to the classical GFP-based systems (Gautier *et al.*, 2008): by using a strong inhibitor, the O^6-benzyl-guanine (O^6-BG), conjugated to fluorescent probes. After the reaction, the probes are covalently bound to the protein. The construct for the expression of the thermophilic variant of the SNAP-tag™ (Visone *et al*., 2017) was obtained substituting the gene 104 from the construct pBS-rRNA-104 with the *S.solfataricus ogt* gene. The structure of the construct termed pBS-rRNA$_p$-OGT is shown schematically in Fig. 4a.

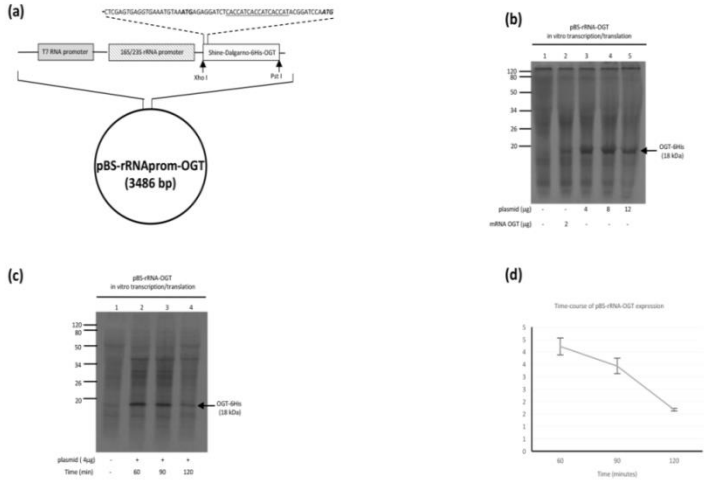

Figure 4: *In vitro* expression of *Ss*OGT. (a) Schematic representation of pBS-rRNA$_p$-*ogt* plasmid. It was designed by introducing a DNA fragment of 522 bp containing the *ogt* gene into the *Xho* I- *Pst* I sites replacing ORF 104. The coding region starts with an AUG codon (bold letters) preceding a DNA region coding for six histidines (underlined letters) placed to the amino-terminal region of the *Ss*OGT protein (bold and italic letters). DNA insert contains a SD motif (italic letters), retained from the ORF 104, located 7 nucleotides upstream from the coding region. (b) Increased amount of pBS-rRNA$_p$-*ogt* plasmid were incubated with *S. solfataricus* whole cell extract for 60 min at 70°C in a final volume of 25 µl and the products of expression were resolved by 16% denaturing polyacrylamide gel electrophoresis. (c) Time course of *Ss*OGT expression: 4 µg of pBS-rRNA$_p$-*ogt* plasmid were incubated with *S. solfataricus* whole cell extract at 70°C and equal aliquotes of the reaction were withdraw from the mixture at the indicated times. (d) Graph is plotted with the values of the band intensity corresponding to *Ss*OGT protein shown in (c) and quantified using ImageJ software (NIH). The values represent the average of three independent experiments. All error bars indicate SD.

Specifically, the strong SD motif 7 nucleotides upstream from the AUG start codon was retained, and preceded His_6- coding triplets followed by the *ogt* open reading frame. As shown in Fig. 4b, the gene was expressed producing a main protein band of about 18 kDa, corresponding to the expected size of the ORF *Ss*OGT-His_6. As a positive control, we employed an *ogt* mRNA transcribed *in vitro* from the T7 promoter (lane 2), which, as expected, was translated less efficiently than the mRNA directly transcribed in the reaction mix. This is possibly due to the different 5'-UTR of the two mRNAs, but it is also conceivable that when translation takesplace at the same time as transcription the mRNA is stabilized and the ribosomes may bind more easily to the translation start sites. To gain insight into other factors influencingthe efficiency of *Ss*OGT protein expression, we analysed the time course of the reaction with fixed amount of the same construct. The highest expression level of the protein was observed after 60 min incubation, while at longer times (90 and 120 min) the efficiency decreased (Fig. 4c and d), as observed in other *in vitro* expression systems (Spirin *et al*., 1988). Furthermore, we tested whether the linearization of the construct could produce a transcriptional run-off at the end of the gene with a consequent increase of the product of our interest. This was not the case, however: samples incubated with the linearized plasmid failed to yield a band corresponding to the expected size of the ORF *Ss*OGT-His_6 (Fig 5a). Further analysis revealed that this was due to degradation of the linearized plasmid in the reaction mix (as illustrated in Fig. 5b) similarly to results obtained by other authors with different cell-free coupled transcription-translation systems (Yang *et al*., 1980).

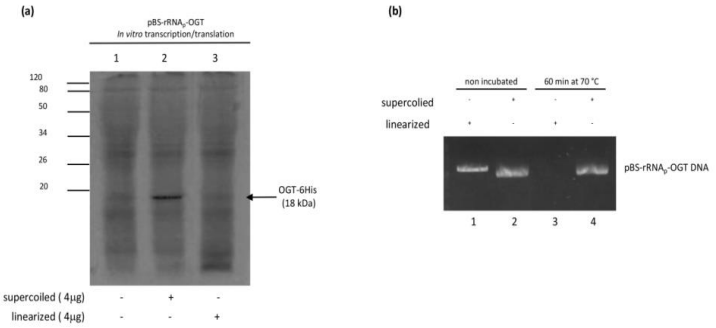

Figure 5: ***In vitro* expression of *ogt* from linearized plasmid.** (a) Supercoiled and linear pBS-rRNA$_p$-*ogt* plasmids were incubated with *S. solfataricus* whole cell extract for 60 min at 70 °C with ^{35}S-Met in a final volume of 25 µl and the products of expression were resolved by 16% denaturing polyacrylamide gel electrophoresis. (b) Survival of supercoiled and linear pBS-rRNA$_p$-*ogt* plasmid after incubation in the S-30 coupled system. The constructs were incubated for 60 min at 70 °C under standard conditions and then analysed on a 1% agarose gel Lanes: 1, non-incubated linear pBS-rRNA$_p$-*ogt* DNA; 2, non-incubated supercoiled pBS-rRNA$_p$-*ogt* DNA; 3, linear pBS-rRNA$_p$-*ogt* DNA incubated in an S-30 mixture; 4, supercoiled pBS-rRNA$_p$-*ogt* DNA incubated in an S-30.

3.4 Characterization of *Ss*OGT activity

To test whether the *in vitro* produced *Ss*OGT was functionally active, we incubated the construct pBS-rRNA$_p$-OGT with the lysate at 70 ° C for 1h in presence of fluorescein-derivated O^6-BG (SNAP-Vista Green™, New England Biolabs). As above mentioned, *Ss*OGT catalyzes the formation of a covalent bond between the benzyl group of BG and a specific cysteine residue in its active site; in our case, by using a fluorescein-derivative of BG, the successful completion of the reaction renders the protein fluorescent. Indeed, we observed a fluorescent band corresponding to the expected size of the *Ss*OGT in the reaction conditions adopted (as shown in Fig. 6), demonstrating the active state of the expressed protein.

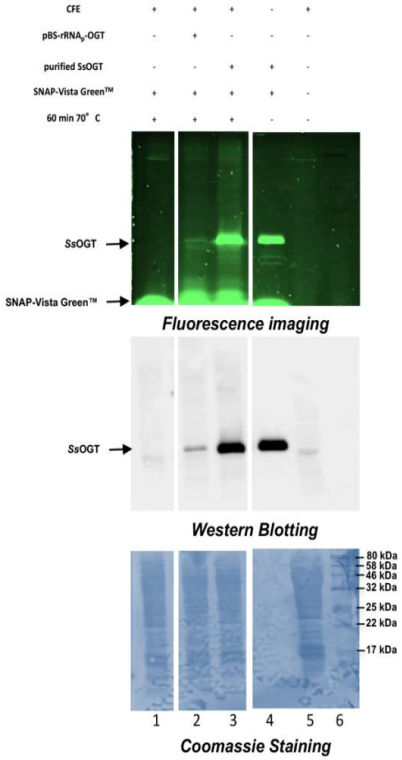

Figure 6. SsOGT labeling. SDS-PAGE of *in vitro* expressed pBS-rRNA$_p$-*ogt* plasmid and purified *Ss*OGT protein both incubated with the BG-FL substrate (5 µm) for 60 min at 70°C. The gel was exposed for fluorescence *gel-imaging* analysis, blotted and stained with Coomassie blue. The filter was probed with the anti-*Ss*OGT antibody(middle panel). Lane 1 contains 100 µg of *S. solfataricus* S30 fraction in presence of the BG-FL substrate; lane 2 contains 8 µg of pBS-rRNA$_p$-*ogt* plasmid in 100 µg of *S. solfataricus* S30 fraction and BG-FL substrate; lane 3 contains 200 ng of purified *Ss*OGT protein with 100 µg of *S.solfataricus* S30 fraction and BG-FL substrate; lane 4 contains 200 ng of purified OGT protein with BG-FL substrate; lane 5 contains 100 µg of *S.solfataricus* S30 fraction; lane 6 corresponds to the protein marker.

The levels of *in vitro* expressed *Ss*OGT were assessed by comparing its fluorescence with that obtained with known amounts of recombinant protein. The outcome of

the experiment permitted also to exclude the possibility that *in vitro* produced *Ss*OGT was degraded after its translation and upon the irreversible transfer of the fluoresceinated-benzyl group to the active site, as previously demonstrated (Perugino *et al*., 2015).In effect, incubation for 60 min at 70 °C of the recombinant *Ss*OGT in the *S. solfataricus* lysates in the presence of the SNAP-Vista Green™ did not affect the activity nor the fluorescent signal obtained (as reported in Fig. 6, lane 3).This analysis allowed us to estimate the amount of *in vitro* translated *Ss*OGT to an order of magnitude, corresponding to ca. 10-20 ng of protein produced for μg of plasmid used, in 25 μl of reaction.

4.DISCUSSION

4.1 Advantages and applications of the *in vitro* transcription/translation system

The present study reports the development of a transcription/translation system for the synthesis of proteins at high temperature (70°C), based on an S30 extractfrom the thermophilic crenarcheon *S. solfataricus*. The system makes use of an engineered classical pBS-SK plasmid, where efficient transcription is driven by a strong promoter, corresponding to the DNA region upstream from the 16S/23S rDNA gene, while translation is stimulated by the presence of a strong SD-motif ahead of the start codon of the chosen gene. The reaction works at the optimal temperature of 70°C and maximal protein synthesis is achieved after 1 h of incubation. A preliminary assessment of the various parameters and components that affect the rate and yield of protein synthesis was performed. We tested the system with two different genes, one encoding a ribosomal protein and another encoding *Ss*OGT, an enzyme, whose activity was determined by using a fluorescent-based assay, as described above. The former gene had already shown to be efficiently translated *in vitro* from a pre-trascribed mRNA (Condò *et al*., 1999) and served as a starting point to tune the system. Transcription/translation of the *ogt*-encoding gene led to an active protein, thereby demonstrating that it was correctly folded/modified in the *in vitro* reaction. Moreover, the possibility to use fluorescent substrates for this enzyme is a clear advantage for the quantification of the gene product, making this system flexible. The simplicity of the experimental procedure and specific activity of the proteins offer a number of possibilities for the study of structure-function relationships of proteins. In addition to the therapeutic and analytical approaches, a highly

investigated field is the industrial large-scale cell-free production of proteins. For istance, (wheat germand *E.coli* systems are already used as cell-free production platforms for vaccines and new therapeutics against malaria (Arumugam *et al*., 2014) and human parainfluenza virus type 3 (Senchi*et al*., 2013) as well as for cytokines and antibodies (Zawada *et al*., 2011; Zimmerman *et al*., 2014). Therefore, we believe that our system will be appropriate to a broad range of applications for basic and applied research. Moreover, the efficient production and characterization of proteins that are difficult to express in living cells (e.g., toxic proteins, several membrane proteins, some post-translationally modified proteins) might provide novel functional and pharmacological insights.

An important novelty of our system with respect to previous attempts described in the literature is the utilization of only endogenous components present in the cell lysate. To date, the *Thermococcus kodakaraensis* lysate is the only described system for protein synthesis coupled with high-temperature translation. However, it requires an added thermostable T7 RNA polymerase. (Endoh *et al*., 2006). Our assay is therefore an economically convenient alternative, since extract preparation is simple and inexpensive.While the present work describes a promising new technology mainly for the gene expression analysis, it is not yet usable as such for the *in vitro* scale-up production of recombinant proteins. To achieve this, further experiments and improvements are needed, for instance, by dividing the reaction in two compartments, one containing the modified extract and one containing a feeding solution that includes substrates such as amino acids, ATP and GTP, and that is renewed by continuous flow, permitting

substrate replenishment and by product removal. Moreover, it should be observed that extant-coupled CFPS utilize DNA in three forms: linear PCR product,linearized plasmid and circular plasmid. The use of linear PCR products has the distinct advantage of simplicity, since it eliminates theneed for time-consuming cloning steps. However, circular DNA plasmidshave typically been preferred to linearized plasmids or PCR products, due to the greater susceptibility of linear DNAs to nucleolytic cleavage. Indeed, in our case, samples incubated with the linearized plasmid failed to yield the expected protein product due to degradation of the linearized plasmid in the reaction mix. The removal of nucleases, and/or the utilization of overhang extensions to cyclize PCR products, is another objective for the future optimization of the system.

In conclusion, we believe that the system described here has very good potential for use in fields such as protein display technologies, interactome analysis and understanding of the molecular mechanisms governing coupled transcription-translation in Archaea.

5.REFERENCES

Akanuma G., Kobayashi A., Suzuki S., Kawamura F., Shiwa Y., Watanabe S., Yoshikawa H., Hanai R. and Ishizuka M. (2014).Defect in the formation of 70S ribosomes caused by lack of ribosomal protein L34 can be suppressed by magnesium.*J Bacteriol.* **196**(22):3820-30.

Andreotti G., Cubellis M.V., Nitti G., Sannia G., Mai X., Adams M.W. and Marino G. (1995). An extremely thermostable aromatic aminotransferase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochim Biophys Acta.* **1247**(1):90–96.

Arumugam T.U., Ito D.,Takashima E.,Tachibana M., Ishino T.,Torii M.and Tsuboi T.(2014). Application of wheat germ cell-free protein expression system for novel malaria vaccine candidate discovery. *Expert Rev Vaccines.* **13**(1):75-85.

Aoki M., Matsuda T., Tomo Y., Miyata Y., Inoue M., Kigawa T. and Yokoyama S. (2009). Automated system for high-throughput protein production using the dialysis cell-free method. *Protein Expr Purif.* **68**(2):128-36.

Bank A. and Marks P.A. (1966). Protein synthesis in a cell-free human reticulocyte system: ribosome function in thalassemia. *J Clin Invest*. **45**(3):330-6.

Benelli D. and Londei P. (2007). *In vitro* studies of Archaeal translational initiation. *Methods Enzymol.* **430**:79-109.

Carlson E.D.,Gan R.,Hodgman C.E. and Jewett M.C. (2012). Cell-free protein synthesis: applications come of age. *Biotechnol Adv.* **30**(5):1185-94.

Chen H.Z. and Zubay G. (1983). Analysis of ColE1 expression *in vitro* after chromosome fragmentation. *J Bacteriol*. **154**(2):650-5.

Condò I., Ciammaruconi A., Benelli D., Ruggero D. and Londei P. (1999). Cis-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. *MolMicrobiol.* **34**(2):377-84.

Craig D., Howell M.T., Gibbs C.L., Hunt T. and Jackson R.J. (1992). Plasmid cDNA-directed protein synthesis in a coupled eukaryotic *in vitro* transcription-translation system. *Nucleic Acids Res.* **20**(19):4987-95.

Endoh T., Kanai T., Sato Y.T., Liu D.V., Yoshikawa K., Atomi H. and Imanaka T.(2006). Cell-free protein synthesis at high temperatures using the lysate of a hyperthermophile. *J Biotechnol.***126**(2): 186-95.

Ezure T., Suzuki T., Higashideet S., Shintani E.,Endo K.,Kobayashi S.,Shikata M., Ito M., Tanimizu K. and Nishimura O. (2006). Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol. Prog*. **22**(6):1570–1577.

Gale E.F. and Folkes J.P.(1954). Effect of nucleic acids on protein synthesis and amino acid incorporation in disrupted staphylococcal cells. *Nature.***173**(4417):1223-7.

Gautier A., Jiullerat A., Heinis C., Reis Corrêa I., Kindermann M., Beaufils F. and Johnsson K. (2008). An engineered protein tag for multiprotein labeling in living cells. *Chemistry & Biology*. **15**(2):128-136.

Jackson R.J., Campbell E.A., Herbert P. and Hunt T. (1983). The preparation and properties of gel filtered rabbit reticulocyte lysate protein synthesis systems. *Eur J Biochem.* **131**(2):289-301.

Josephsen J. and Gaastra W. (1985). DNA directed *in vitro* protein synthesis with *Escherichia coli* S-30 extracts. *DNA Methods Mol Biol.* **2**:131-5.

Keller E.B. and Zamecnik P.C. (1956). The effect of guanosine diphosphate and triphosphate on the incorporation of labeled amino acids into proteins. *J Biol Chem.* **221**(1):45-59.

Keller E.B and Littlefield J.W. (1957). Incorporation of C14-amino acids into ribonucleoprotein particles from the Ehrlich mouse ascites tumor. *J Biol Chem*. **224**(1):13-30.

Krieg P.A. and Melton D.A. (1987). *In vitro* RNA synthesis with SP6 RNA polymerase. *Methods Enzymol.* **155**:397-415.

Lamborg M.R. and Zamecnik P.C. (1960). Amino acid incorporation into protein by extracts of *E. coli*. *BiochimBiophysActa*. **12**(42): 206–211.

Londei P., Teixidò J., Acca M., Cammarano P. and Amils R. (1986). Total reconstitution of active large ribosomal subunits of the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. *Nucleic Acids Res.* **14**(5):2269-85.

Marcus A. and Feeley J. (1966). Ribosome activation and polysome formation *in vitro*: requirement for ATP. *Proc Natl AcadSci U S A.* **56**(6):1770-7.

McQuillen K., Roberts R.B. and Britten R.J. (1959). Synthesis of nascent protein by ribosomes in *Escherichia coli*. *Proc Natl Acad Sci U S A*. **45**(9):1437-47.

Nierhaus K.H. (2014). Mg^{2+}, $K^{+}$and the ribosome. *J Bacteriol.* **196**(22):3817-9.

Niremberg M.W. and Matthaei J.H. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. USA* **47**:1558-1602.

Pelham H.R. and Jackson R.J. (1976). An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem.* **67**(1): 247–256.

Perugino G., Miggiano R., Serpe M., Vettone A., Valenti A., Lahiri S., Rossi F., Rossi M., Rizzi M. and Ciaramella M. (2015). Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein.*Nucleic Acids Res.* **43**(18): 8801-16.

Perugino G., Vettone A., Illiano G., Valenti A., Ferrara M.C., Rossi M. and Ciaramella M. (2012). Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage.*J Biol Chem*. **287**(6):4222-31.

Petermann M.L. and Hamilton M.G.(1957). The purification and properties of cytoplasmic ribonucleoprotein from rat liver. *J Biol Chem*. **224**(2):725-36.

Qureshi S.A.,Baumann P.,Rowlands T.,Khoo B. and Jackson S.P. (1995). Cloning and functional analysis of the TATA binding protein from *Sulfolobus shibatae*. *Nucleic Acids Res*. **23**(10):1775-81.

Reiter W.D., Hüdepohl U. and Zillig W.(1990). Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency

and start-site selection *in vitro*. *Proc Natl Acad Sci USA.* **87**(24):9509-13.

Ruggero D., Creti R. and Londei P. (1993). *In vitro* translation of Archaeal natural mRNAs at high temperature. *FEMS Microbiol Lett.* **107**: 89–94.

Schweet R.S., Bovard F.C., Allen E. and Glassman E. (1958). The incorporation of amino acids into ribonucleic acid. *Proc Natl Acad Sci U S A.* **44**(2):173-7.

Senchi K., Matsunaga S., Hasegawa H., Kimura H. and Ryo A. (2013). Development of oligomannose-coated liposome-based nasal vaccine against human parainfluenza virus type 3. *Front Microbiol.* **4**:346.

Shimizu Y., Inoue A., Tomari Y., Suzuki T., Yokogawa T., Nishikawa K. and Ueda T. (2001). Cell-free translation reconstituted with purified components. *NatBiotechnol.* **19**(8):751-5.

Shirokov V.A., Kommer A., Kolb V.A. and Spirin A.S. (2007). Continuous-exchange protein-synthesizing systems. *Methods Mol Biol.***375**:19-55.

Spirin A.S., Baranov V.I., Ryabova L.A., Ovodov S.Y. and Alakhov Y.B.(1988). A continuous cell-free translation system capable of producing polypeptides in high yield. *Science*. **242**:1162–1164.

Swartz J. (2006). Developing cell-free biology for industrial applications. *J IndMicrobiolBiotechnol.* **33**(7):476-85.

Takai K., Sawasaki T. and Endo Y. (2010). Practical cell-free protein synthesis system using purified wheat embryos. *Nat Protoc*. **5**(2):227-38.

Vinarov D.A. and MarkleyJ.L. (2005). High-throughput automated platform for nuclear magnetic resonance-based structural proteomics. *Expert Rev Proteomics.* **2**(1):49-55.

Visone V., Han W., Perugino G., Del Monaco G., She Q., Rossi M., Valenti A. and Ciaramella M. (2017). *In vivo* and *in vitro* protein imaging in thermophilic Archaea by exploiting a novel protein tag. *PLoS One* **12**(10):e0185791.

Ward D.E., Kengen S.V., van Der Oost J. and de Vos V.M. (2000). Purification and characterization of the alanine aminotransferase from the hyperthermophilic archaeon *pyrococcus furiosus* and its role in alanine production. *J Bacteriol.* **182**(9):2559-66.

Watanabe M., Miyazono K., Tanokura M.Sawasaki T., Endo Y. and Kobayashi I. (2010). Cell-free protein synthesis for structure determination by X-ray crystallography. *Methods Mol. Biol.* **607**: 149–160.

Yang H.L., Ivashkiv L., Chen H.Z., Zubay G. and Cashel M.(1980). Cell-free coupled transcription-translation system for investigation of linear DNA segments. *Proc Natl Acad Sci U S A.* **77**(12): 7029-33.

Yang J., Kanter G., Voloshin A., Michel-Reydellet N., Velkeen H., Levy R. and Swartz J.R. (2005). Rapid expression of vaccine proteins for B-cell lymphoma in a cell-free system. *BiotechnolBioeng.***89**(5):503-11.

Zamecnick P.C. (1969). An historical account of protein synthesis, with current overtones a personalized view. *Cold Spring HarbSymp Quant Biol.* **34**:1-16.

Zawada J.F., Yin G., Steiner A.R., Yang J., Naresh A., Roy S.M., Gold D.S., Heinsohn H.G. and Murray C.J. (2011). Microscale to manufacturing scale-up of cell-free cytokine production a new approach for shortening protein production development timelines. *BiotechnolBioeng.* **108**(7):1570-8.

Zimmerman E.S**.**, Heibeck T.H., Gill A., Li X., Murray C.J., Madlansacay M.R., Tran C., Uter N.T., Yin G., Rivers P.J., Yam A.Y., Wang W.D., Steiner A.R., Bajad S.U., Penta K., Yang W., Hallam T.J., Thanos C.D. and Sato A.K. (2014). Production of site-specific antibody-drug conjugates using optimized non-natural amino acids in a cell-free expression system. *Bioconjug Chem.* **25**(2)**:**351-61.

Zubay G. (1973). *In vitro* synthesis of protein in microbial systems. *Annu Rev Genet.* **7**:267-287.