Assembling the archaeal ribosome: roles for translation-factor-related GTPases

Fabian Blombach1, Stan J.J. Brouns and John van der Oost
Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands

Abstract
The assembly of ribosomal subunits from their individual components (rRNA and ribosomal proteins) requires the assistance of a multitude of factors in order to control and increase the efficiency of the assembly process. GTPases of the TRAFAC (translation-factor-related) class constitute a major type of ribosome-assembly factor in Eukaryota and Bacteria. They are thought to aid the stepwise assembly of ribosomal subunits through a ‘molecular switch’ mechanism that involves conformational changes in response to GTP hydrolysis. Most conserved TRAFAC GTPases are involved in ribosome assembly or other translation-associated processes. They typically interact with ribosomal subunits, but in many cases, the exact role that these GTPases play remains unclear. Previous studies almost exclusively focused on the systems of Bacteria and Eukaryota. Archaea possess several conserved TRAFAC GTPases as well, with some GTPase families being present only in the archaeo-eukaryotic lineage. In the present paper, we review the occurrence of TRAFAC GTPases with translation-associated functions in Archaea.

Introduction
Translation of the genetic information from nucleic acids into proteins is a highly conserved process in all three domains of life (Eukaryota, Bacteria and Archaea). The key player in this process is the ribosome, a large ribonucleoprotein complex composed of three or four distinct rRNAs and 55–87 ribosomal proteins [1]. The ribosome catalyses the formation of peptide bonds during protein synthesis, a process that requires the assistance of translation factors to achieve precision and efficiency during the initiation, elongation and termination steps of translation. Some translation factors are universal, i.e. found in all three domains of life. Archaea employ translation factors that are a subset of the characterized eukaryotic translation factors, whereas only the universally conserved factors are shared with Bacteria [2].

Several translation factors are GTPases, constituting the class of so-called TRAFAC (translation-factor-related) GTPases [3]. This tight functional association of TRAFAC GTPases with the translation machinery suggests that they have co-evolved with the translation machinery. In fact, TRAFAC GTPases are also involved in other translation-associated processes such as tRNA modification and ribosome assembly [3]. Apart from the aforementioned translation factors, a few TRAFAC GTPase families are considered to be universal, whereas several other GTPase families are widely distributed in two of the three domains [4]. Little is known about the function of most of these highly conserved GTPases [5].

Key words: archaeon, GTPase, ribosome assembly, translation-factor-related GTPase (TRAFAC GTPase).

Abbreviations used: a, archaeal; e, eukaryotic; EF, elongation factor; IF, initiation factor; LSU, large ribosomal subunit; TRAFAC, translation-factor-related; YRG, YlqF-related GTPase.

1To whom correspondence should be addressed (email fabian.blombach@gmail.com).

G-domain
The common denominator of GTPases is the highly conserved G-domain that is responsible for binding and hydrolysis of guanine nucleotides. GTPases are considered to work as ‘molecular switches’: they undergo conformational changes when switching between the GTP- and the GDP-bound form, corresponding to the ‘on’- and ‘off’-state of the GTPase respectively [6]. The on- and off-states of the GTPase are transduced into different structural conformations of the two mobile regions of the G-domain, termed switch I and switch II.

The majority of ribosome-associated TRAFAC GTPases are multidomain proteins in which the additional domains are involved in ligand interactions. Curiously, although ribosomal subunits are the common ligand for these GTPases, the additional domains that mediate the interaction are non-orthologous. In some cases, these are well-known RNA-binding domains such as KH (K homology) domains (e.g. [7]), whereas in other cases, they are unique domains whose occurrence is restricted to this GTPase family.

Ribosome assembly in Archaea: prokaryotic and eukaryotic features
The high number of ribosomal proteins that need to be incorporated into the ribosome requires the assembly process to be organized in a stepwise manner. In both Bacteria and Eukaryota, GTPases function in the organization of the stepwise assembly by serving as checkpoints that regulate the recruitment of additional ribosome-assembly factors or ribosomal proteins to ribosome precursors [8]. The compartmentalization of the eukaryotic cell with separate locations for the transcription of rRNA and the translation of ribosomal proteins poses a problem for ribosome assembly. Ribosomal proteins have to be imported into the nucleolus in...
order to make use of the ‘assembly gradient’, the binding of
the first ribosomal proteins to the nascent rRNA transcript,
thereby establishing a 5′→3′ order of assembly. Subsequently,
the ribosomal subunit precursors assembled in the nucleolus
have to be exported into the cytoplasm, a process that
proceeds via two different routes for the LSU (large ribosomal
subunit) and the SSU (small ribosomal subunit).

More than 200 proteins identified to date are involved in
eukaryotic ribosome assembly [9,10]. Only a small minority
of those factors show clear orthology to archaeal genes,
including the RNA-guided machineries carrying out the
modifications of rRNA nucleotides [11]. Archaea can be
predicted to employ fewer factors for ribosome assembly
compared with Eukaryota, and it will be interesting to
see to what extent the archaeal ribosome assembly process
resembles the corresponding process in Bacteria [12].

The classical translation factors and SelB
The GTPase family of classical translation factors contains
the four universally conserved translation factor subfamilies
with their archaeal representatives EF (elongation factor) 1α,
EF2, e/aIF (eukaryotic/archaeal initiation factor) 2γ/SelB and
aIF2/SB (Table 1). Their functions in translation initiation
and elongation are relatively well characterized.

Archaeal orthologues of the universal selenocysteinespecific
translation elongation factor SelB are widely distributed
in Crenarcheota, Thaumarcheota and Euryarcheota,
but are not strictly conserved (Table 1). It functions analogous
to EF1α/EF-Tu by delivering during translation elongation
e a selenocysteine-charged tRNA to internal UGA codons
present in certain mRNAs. UGA normally serves as one
of the three stop codons. SelB orthologues from the
Methanococcales have been studied in more detail. Gene
disruption confirmed their involvement in selenoprotein
synthesis [13]. The structure of the SelB orthologue from
*Methanococcus maripaludis* revealed structural homology
with aIF2 and bacterial IF2 [14]. Domain IV might be
involved in the binding to the mRNA to receive the recoding
signal [14]. Interestingly, domain IV is conserved only in
few SelB orthologues from methanogenic Archaea, but it
is generally absent from other archaeal SelB orthologues.
In addition, there is no strict co-occurrence of other
components such as the selenocysteine-specific tRNA with
SelB orthologues, implying that there is some variation in
the mechanism of selenocysteine incorporation in the different
organisms or alternatively that some SelB GTPases might
have a different function.

GTPases predicted to function in ribosome assembly

The YRG (YlqF-related GTPase) family and its
archaeal member MJ1464
A remarkable feature of the YRG family is the circular
permutation of the G-domain. Various members of the
different subfamilies of YRG have been shown to participate
in bacterial and eukaryotic ribosome assembly [15–21].
Bacteria and Eukaryota possess several paralogous members
of the YRG family. The MJ1464 subfamily encompasses the
archaeal members of the YRG family. MJ1464 GTPases
are present in all Crenarchaeota and *Korarchaeum*, and
in some Euryarchaeota, but (at present) they seem absent from
the Thaumarchaeota (Table 1).

*Bacillus subtilis* YlqF/RbgA loads the ribosomal protein
L16 on to a 45S precursor of the LSU [21]. Similarly,
the cytoplasmic GTPase Lsg1p from yeast is required for
the loading of the ribosomal protein L10e on to the LSU
as well as for the release of the archaeo-eukaryotic ‘ribosome
export factor’ Nmd3 [15,16]. Bacterial L16 and eukaryotic
L10e are orthologous proteins, suggesting that YlqF/RbgA
and Lsg1 might share an evolutionarily conserved function
of YRG GTPases. They might play a similar role in what can
be seen as a quality check of the LSU structure [8]. It will be
interesting to see whether MJ1464 is involved in loading of
L10e on to the LSU in Archaea.

The archaeal MJ1464 GTPases encompass a C-terminal
domain that is homologous with the C-terminal domains
of bacterial YlqF GTPases [22] and the eukaryotic ribosome-
assembly GTPase Nug1 [17]. This underlines the potential
role of MJ1464 GTPases in ribosome assembly.

YihA/EngB/YsxC
Members of the YihA/EngB/YsxC GTPase family are
ubiquitous in Eukaryota and Bacteria, but in Archaea, their
occurrence is restricted to the euryarchaeal branch (Table 1).
The structure of the YsxC orthologue from *Pyrococcus
horikoshii* PH0200 in the GDP-bound state has been solved
by the RIKEN Structural Genomics Initiative (PDB code
2CXX). Comparison with the published GDP-bound YsxC
structure from *B. subtilis* [23] (PDB code 1SVI) reveals
overall good structural conservation [rmsd (root mean square
deviation) of 1.9 Å (1 Å = 0.1 nm)] with two additional
α-helices being present in PH0200 (Figure 1). Whereas
canonical G-domains from the TRAFAC class have a central
six-stranded β-sheet, the β-sheet of the G-domain of *B.
subtilis* YsxC contains an additional N-terminal β-strand.
In contrast, the archaeal YsxC orthologues possess a canonical
six-stranded β-sheet. A positively charged patch on the
surface of YsxC consisting of Arg31, Arg116, His117 and Lys146
[23] was proposed to be involved in the interaction of YsxC
with the LSU [24]. In PH200, this patch is partially conserved
(Arg37, Lys119 and Lys162), which might indicate functional
conservation.

NOG1
NOG1 GTPases are part of the Obg family of GTPases [3]
and are highly conserved in Eukaryota and Archaea except for
Thaumarchaeota (Table 1). Eukaryotic NOG1 is a nucleolar
GTPase involved in assembly of the LSU [16,25]. Yeast
NOG1 interacts directly with the ribosome assembly protein
Rlp24, a eukaryotic parologue of the ribosomal protein L24e
[26]. Archaea have only a single L24e orthologue that is part of
mature LSUs in *Haloarcula marismortui* [27]. Interestingly,
Table 1 | TRAFAC GTPases with characterized or predicted roles in translation

Accession numbers are presented for the GTPases.

<table>
<thead>
<tr>
<th>GTPase</th>
<th>Thaumarchaeota</th>
<th>Crenarchaeota</th>
<th>Euryarchaeota</th>
<th>Korarchaeota</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nitrosopumilus maritimus</td>
<td>Sulfolobus solfataricus</td>
<td>Pyrobaculum aerophilum</td>
<td>Pyrococcus horikoshii</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The classical translation factor family</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α/εIF2γ</td>
<td>Nmar_0072</td>
<td>SSO0412</td>
<td>PAE0064</td>
<td>PH1706</td>
</tr>
<tr>
<td>αIF2/3B</td>
<td>Nmar_0229</td>
<td>SSO0228</td>
<td>PAE1513</td>
<td>PH1095</td>
</tr>
<tr>
<td>EF1α</td>
<td>Nmar_1034</td>
<td>SSO0216</td>
<td>PAE3000</td>
<td>PH1484</td>
</tr>
<tr>
<td>EF2</td>
<td>Nmar_0748</td>
<td>SSO0728</td>
<td>PAE0332</td>
<td>PH1899</td>
</tr>
<tr>
<td>SelB</td>
<td>Nmar_0961</td>
<td>SSO0470</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GTPases predicted to function in ribosome assembly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ1464</td>
<td>–</td>
<td>SSO0581</td>
<td>PAE0759</td>
<td>PH0645</td>
</tr>
<tr>
<td>YssX</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>PH0200</td>
</tr>
<tr>
<td>NOG1</td>
<td>–</td>
<td>SSO0625</td>
<td>PAE3626</td>
<td>PH1320</td>
</tr>
<tr>
<td>Ribosome-binding GTPases of unknown function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRG</td>
<td>Nmar_0684</td>
<td>SSO2385</td>
<td>PAE1649</td>
<td>PH1672</td>
</tr>
<tr>
<td>Ygi210</td>
<td>Nmar_1534</td>
<td>SSO0743</td>
<td>PAE3165</td>
<td>PH0525</td>
</tr>
<tr>
<td>HIX</td>
<td>Nmar_1605</td>
<td>SSO0269</td>
<td>PAE0782</td>
<td>PH1132</td>
</tr>
<tr>
<td>GP1</td>
<td>–</td>
<td>SSO2429</td>
<td>PAE1764</td>
<td>–</td>
</tr>
<tr>
<td>Other archaeal proteins with potential roles in ribosome assembly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nmd3</td>
<td>–</td>
<td>SSO2383</td>
<td>PAE1653</td>
<td>PH0990</td>
</tr>
<tr>
<td>SBDS</td>
<td>Nmar_0430</td>
<td>SSO0737</td>
<td>PAE2209</td>
<td>PH1552</td>
</tr>
<tr>
<td>αIF6</td>
<td>Nmar_0237</td>
<td>SSO0351</td>
<td>PAE097</td>
<td>PH0528</td>
</tr>
</tbody>
</table>
Figure 1 | Comparison between (A) YsxC from *Bacillus subtilis* in GDP-bound form (PDB code 1SVI) and (B) the YsxC orthologue PH200 from *P. horikoshii* in GDP-bound form (PDB code 2CXX)

\(\alpha\)-Helices and \(\beta\)-strands present only in one of the two orthologous proteins are coloured green. The GDP molecule is shown in blue. Residues forming a conserved positively charged surface patch are shown in red.

the archaeal L24e orthologue shares specific features with Rlp24 such as two pairs of cysteine residues [26]. Thus it seems possible that the archaeal L24e has a function in ribosome assembly as well and might interact with the archaeal NOG1 orthologue. This would be similar to ribosomal protein L7ae that has a double life as a structural component of the ribosome as well as a ribosome-assembly factor in the RNA-guided rRNA-modification machinery [28].

The crystal structure of the NOG1 orthologue PH1320 from *P. horikoshii* in its GDP-bound form has been solved by the RIKEN Structural Genomics Initiative (PDB code 2E87). The N-terminal 160 amino acids comprise a four-helical bundle that provides positively charged surface patches that might be involved in ribosome binding (Figure 2). The two mobile elements, switch I and switch II, of the G-domain mediate contacts with the N-terminal domain following a pattern commonly found in multidomain GTPases for the interaction of the G-domain with a ligand-binding domain [7,29–31].

**Ribosome-binding GTPases of unknown function**

Apart from the classical translation factors and GTPases where a role in ribosome assembly has been demonstrated, several other TRAFAC GTPases interact with ribosomal subunits, but their biological role remains to be determined. For other TRAFAC GTPases, no functional studies are available, but on the basis of their domain composition, a translation-related function can be predicted.

**Obg family GTPases DRG and Ygr210**

Next to NOG1, two more GTPases of the Obg family are present in Archaea: DRG and Ygr210 (Table 1) [3]. In both DRG and Ygr210, an N-terminal G-domain is followed by a C-terminal TGS (ThrRS, GTPases, SpoT) domain [3]. This domain is also present in other Obg family GTPases, and its structure in *B. subtilis* Obg [29] and *Haemophilus influenzae* YchF has been characterized [32]. Besides Obg family GTPases, TGS domains have also been found in threonyl-tRNA synthetases (ThrRS) and guanosine polyphosphatases (SpoT) [33].

Several bacterial Obg GTPases have been shown to associate with free 50S ribosomal subunits [34–36]. Obg GTPases control stringent response in Bacteria, but, in addition, they might play a role in ribosome assembly [34,35,37]. YchF is still largely uncharacterized, but the structure of *H. influenzae* YchF suggests binding of double-stranded nucleic acids [32]. *Trypanosoma cruzi* YchF possibly interacts with the translation machinery [38]. On the basis of this circumstantial evidence, it is likely that DRG and Ygr210 GTPases have translation-related functions as well. The G-domains of the Ygr210 family have canonical G4 sequence motifs conferring specificity for guanine nucleotides, in contrast with the closely related YchF subfamily in which mutated G4 motifs turn them into ATPases [39].

**HflX**

The HflX GTPase family is related to Obg GTPases [3]. HflX GTPases are ubiquitous in all three domains of life, but several taxonomic groups do not contain HflX orthologues. In Archaea, HflX GTPases are present in Crenarchaeota and in the available thaumarchaeal genomes, but are absent from some euryarchaeal species (Table 1). The HflX family is the only TRAFAC GTPase family, apart from the classical translation factors, where experimental data about ribosome interaction is available from an archaeal representative. The
Figure 2 | Structure of the P. horikoshii orthologue PH1320 of the NOG1 GTPase family (PDB code 2E87)

(A) Two-domain arrangement of Nog1 with an N-terminal G-domain (cyan) and a C-terminal domain forming a four-helical bundle (green). The switch I (sw I) and switch II (sw II) regions are highlighted in red and blue respectively. (B) The electrostatic surface representation on the solvent-accessible surface of PH1320 (−10 kT/e to 10 kT/e). The Figure was generated using PyMOL (DeLano Scientific; http://www.pymol.org) and the APBS package [46].

structure of the HflX orthologue from Sulfolobus solfataricus has been solved in apo- and GDP-bound form, revealing the presence of a novel putative RNA-binding domain termed the HflX domain [40]. Similar to the bacterial HflX orthologues from Escherichia coli and Chlamydomphila pneumoniae [41], the archaeal HflX orthologue binds to the LSU in both the GTP- and GDP-bound form (F. Blombach, H. Launay, V. Zorraquino, D. Swarts, L. Cabrita, D. Benalli, J. Christodoulou, P. Londei and J. van der Oost, unpublished work).

GP-1, an uncharacterized GTPase of the classical translation factor family

Another GTPase from the classical translation factor family that is widely spread within Archaea is GP-1 [3]. It is found in all Crenarchaeota as well as several eukaryotic and euryarchaeal species, but it is absent from Thaumarchaeota (Table 1). Several Eukaryota possess multiple highly conserved GP-1 paralogues, whereas the protein is absent from some eukaryotic species. No GP-1 GTPase orthologue has been thoroughly characterized, but it has been demonstrated that mouse GP-1 locates to the cytoplasm [42]. GP-1 GTPases comprise an N-terminus that shows no similarity to other proteins at sequence level. It is followed by a canonical G-domain and two domains homologous with domains II and III of EF1α/EF-Tu. The domain composition clearly suggests a function involving ribosome binding as for other GTPases derived from the classical translation factors such as SelB or LepA [43].

Conclusions

Archaeal genomes encode a number of TRAFAC GTPases with translation-related functions. The absence of NOG1, GP-1 and YRG family GTPases from Thaumarchaeota corroborates further the recent re-classification of these ‘marine Crenarchaeota’ as a separate phylum next to Euryarchaeota and Crenarchaeota [44]. The number of TRAFAC GTPases encoded in thaumarchaeal genomes is surprisingly small, with only three TRAFAC GTPases (Ygr210, HflX and DRG) serving as candidates for ribosome assembly GTPases.

The use of GTPases in ribosome assembly might allow the cells to gain control over the process of ribosome assembly [8]. In Bacteria, decreasing GTP levels during the cellular stringent response to stress conditions would couple ribosome assembly directly to the energy state level of the cell. Interestingly, the stringent response in Sulfolobus species appears not to correlate with decreasing GTP levels [45] suggesting that, at least in the Sulfolobus genus, ribosome assembly might not be regulated in this way.

The cellular compartmentalization of eukaryotic cells marks a sharp difference in ribosome assembly between Eukaryota and Archaea. Studies of the archaeal orthologues of eukaryotic ribosome-assembly factors and of their function in a prokaryotic context are likely to reveal more insight into the far more complex eukaryotic ribosome assembly and into the evolution of compartmentalization of the eukaryotic cell.

Funding

This work was supported financially by NWO (Netherlands Organisation for Scientific Research) [ALW-Vici project 865.05.001 to J.v.d.O.].

References
