When DNA replication and protein synthesis come together

Jonathan Berthon1,2, Ryosuke Fujikane1 and Patrick Forterre1

1 Université Paris-Sud 11, CNRS UMR 8621, Institut de Génétique et Microbiologie, 91405 Orsay Cedex, France
2 Laboratory of Protein Chemistry & Engineering, Department of Genetic Resources Technology, Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Fukuoka 812-8581, Japan

In all organisms, DNA and protein are synthesized by dedicated, but unrelated, machineries that move along distinct templates with no apparent coordination. Therefore, connections between DNA replication and translation are a priori unexpected. However, recent findings support the existence of such connections throughout the three domains of life. In particular, we recently identified in archaeal genomes a conserved association between genes encoding DNA replication and ribosome-related proteins which all have eukaryotic homologs. We believe that this gene organization is biologically relevant and, moreover, that it suggests the existence of a mechanism coupling DNA replication and translation in Archaea and Eukarya.

Emerging links between DNA replication and translation

Connections between replication, repair, and transcription have been described for many years and are now well recognized and actively investigated [1]. By contrast, the search for interplay between DNA replication and protein synthesis has received relatively little attention. Here, we review several unexpected results, recently obtained in Archaea, Bacteria and Eukarya, which point to the existence of widespread and possibly ubiquitous mechanisms which couple protein and DNA syntheses. Recent experimental data indicate that bacterial protein synthesis activity might be linked to DNA replication via (p)ppGpp (guanosine polyphosphate derivatives) and/or small GTPases of the Obg (Spo0B-associated GTP-binding protein) family. Unexpectedly, complexes containing ribosomal proteins, factors involved in ribosome biogenesis, and proteins essential for DNA replication initiation have been described in Eukarya. We also focus on our recent description of a conserved association of archaenal genes that encode DNA replication proteins and proteins involved in translation and/or ribosome biogenesis [2]. We contend that this gene cluster points to the existence of a mechanism that couples DNA replication and translation [2]. We believe that this potential interplay between DNA replication and translation could have major repercussions in various realms of cellular biology, thus prompting the need for further investigations.

The stringent response and DNA replication in Bacteria

In Bacteria, amino acid starvation elicits a sharp increase of intracellular (p)ppGpp concentration that triggers a rapid inhibition of rRNA gene transcription and protein synthesis (the so-called stringent response, reviewed in [3]). Wang and colleagues recently showed that activation of the stringent response in Bacillus subtilis also rapidly blocks the elongation step of DNA replication independent of the chromosomal location of the replication forks [4]. This finding was unexpected because previous work suggested that this arrest could occur only within a specific region near the origin and in cooperation with the replication terminator protein [5]. Interestingly, in vitro experiments demonstrate that (p)ppGpp inhibits the B. subtilis DnaG primase, thus suggesting that intracellular (p)ppGpp increases could directly inhibit Okazaki fragment synthesis [4]. This mechanism could prevent the disruption of replication forks that might otherwise occur.

Glossary

Erb1 (Eukaryotic ribosome biogenesis 1): this protein is required for maturation of the eukaryotic 60S ribosomal subunit. This Saccharomyces cerevisiae protein is homologous to the mammalian protein BOP1 (block of proliferation).

GIRNS (Go Ishi Ni San (five, one, two, three in Japanese)): this complex, originally identified in S. cerevisiae and Xenopus laevis, is involved in the establishment and the progression of DNA replication forks in Eukarya and Archaea.

NOG1 (Nuclear G-protein): this protein, which is located in the nucleolus, is required for 60S ribosomal subunit biogenesis.

Obg (Spo0B-associated GTP-binding protein): this protein, originally identified in an analysis of the transcripts of the B. subtilis spo0B stage 0 sporulation operon, is the founding member of a conserved family of GTP-binding proteins which are involved in various cellular processes, including ribosome biogenesis, sporulation, chromosome segregation, DNA replication, and regulation of the stringent response; the bacterial homologs are designated by various names, including CgtA in Caulobacter crescentus, YbhZ in Haemophilus influenzae, or ObgE in Escherichia coli (CgtAE and YbhZ are sometimes used). (p)ppGpp: an acronym used to refer to the guanosine polyphosphate derivatives guanosine 5’-diphosphate 3’-diphosphate (ppGpp) and guanosine 5’-triphosphate 3’-diphosphate (pppGpp) that are synthesized upon nutrient starvation in bacterial organisms. ppGpp and pppGpp are synthesized from ATP and GMP (ppGpp) and ATP and GDP (pppGpp) by the (p)ppGpp synthase RelA. A RelA homolog designated SpaT hydrolyses (p)ppGpp once the amino acid balance is restored within the cell.

RelA: an enzyme that synthesizes (p)ppGpp in response to amino acid starvation.

Rbr1 (Regulator of ribosome biogenesis 1): this nuclear protein participates in the regulation of ribosome biosynthesis in budding yeast. The human homolog is known as GRWD (glutamate-rich WD repeat-containing protein).

SpaT: an enzyme that hydrolyses (p)ppGpp when amino acids are available.

Yph1p: budding yeast pescadillo homolog. This protein (also known as Nop7p) is required for 60S ribosomal subunit biogenesis. The human protein is known as PES1, after pescadillo (the gene originally identified in zebrafish).

0968-0004/$ – see front matter © 2009 Elsevier Ltd. All rights reserved. doi:10.1016/j.tibs.2009.05.004
due to the starvation-induced reduction of the cellular dNTP pool [4].

All bacterial genomes harbor a gene that encodes DnaG. Most also encode the RelA-SpoT homologs (rsh) proteins [6], whose synthase and hydrolase activities regulate intracellular (p)ppGpp levels. Thus, the mechanism coupling the elongation step of DNA replication with protein synthesis via the stringent response might be widespread throughout Bacteria. However, some findings indicate that the stringent response operates differently in other bacteria. Indeed, ppGpp accumulation in Escherichia coli triggers an immediate inhibition of DNA replication initiation, whereas elongation of ongoing DNA replication forks continues to termination [7,8]. Recently, (p)ppGpp accumulation during the stringent response in Caulobacter crescentus was shown to inhibit the initiation step of DNA replication by promoting, via an unknown mechanism, the degradation of the initiator protein DnaA [9]. Therefore, modulation of DNA replication upon nutrient starvation appears to be an important aspect of bacterial physiology, as B. subtilis, E. coli, and C. crescentus are evolutionarily distant lineages. However, the features of this modulation might differ from one bacterial lineage to another.

Most of the proteins involved in the bacterial stringent response and in the related cross-talk between DNA replication and translation, including DnaG, DnaA and RelA-SpoT, lack homologs in Archaea and Eukarya (plants are an exception: RSH proteins are derived from the cyanobacterial ancestor of the chloroplast [10]). However, in considering the importance of DNA replication and protein synthesis in cell physiology (particularly in terms of energy cost), we believe that mechanisms coordinating these two fundamental cellular processes also operate in these two domains of life. One such system could be built upon universal Obg superfamily proteins.

**Universal GTPases might couple DNA and protein syntheses**

Essential and abundant bacterial GTPases named Obg recently emerged as possible regulators to couple ribosome biogenesis and DNA replication. These proteins function in ribosome biogenesis [11], are required for chromosome segregation [12], have a role in the regulation of the stringent response [13], and are suspected to stabilize arrested DNA replication forks [14]. As a way to rationalize the multiple functions of Obg proteins, it has been proposed that these proteins might sense fluctuations in cellular GDP and GTP concentrations and respond accordingly in a coordinated fashion [15]. ObgE, the E. coli Obg family member, is associated with the 50S precursor of the large ribosome subunit in the exponential growth phase, but they dissociate in the stationary phase or during the stringent response [13]. Aside from its role in ribosome biogenesis, ObgE regulates (p)ppGpp levels by promoting SpoT-dependent (p)ppGpp hydrolysis during favorable growth conditions [13]. However, ObgE depletion results primarily in chromosome segregation and cell division defects [12,16]. In addition, obgE mutants display hypersensitivity to DNA replication inhibitors such as hydroxyurea (a specific inhibitor of ribonucleotide reductase), although DNA replication elongation appears to occur normally in these cells [14]. This finding points to a requirement for Obg proteins in preventing DNA replication fork collapse during conditions of low intracellular dNTP levels [14]. Obg proteins are therefore good candidates to link bacterial DNA replication with protein synthesis via its association with the ribosome, the chromosome and/or the replication forks.

Bacterial Obg proteins have several homologs in Eukarya and Archaea that have been classified within different subfamilies [17]. Like the bacterial Obg homologs, members of the NOG1 (Nucleolar G-protein) subfamily, which is specific to Eukarya and Archaea [17], are likely involved in ribosome biogenesis [18]. The yeast homolog, Nog1p, is a component of a complex that contains proteins involved in 60S ribosomal subunit biogenesis, including Erb1p (Eukaryotic ribosome biogenesis; BOP1 in mammals) and several ribosomal proteins [19]. Remarkably, this complex also contains several proteins essential for DNA replication initiation, including Orc6p (one subunit of the Origin Recognition Complex; ORC6L in human), Mcm6p and three other subunits of the MCM complex (the replicative helicase) [19]. Another component of this large complex, Yph1p (Yeast pescadillo homolog; PES1 in human), which is required for ribosome biogenesis, might modulate the rate of cell proliferation in response to energy sources [19]. Erb1p (Regulator of ribosome biogenesis 1; GRWD in human), another protein involved in early ribosome assembly, interacts also with Yph1p and Orc6p [20]. Inactivation of RRBI alters chromosome segregation, with no apparent DNA replication defect [20]. In human cells, transient depletion of GRWD, PES1, or ORC6L increases the occurrence of abnormal mitoses [20]. Killian and co-workers thus proposed that alterations in proteins that link ribosome biogenesis and DNA replication might directly cause chromosome instability and tumor formation [20]. Taken together, these data strongly suggest the existence of a large protein network, involving at least one Obg homolog, that connects ribosome biogenesis, DNA replication, and chromosome segregation in Eukarya. However, the precise molecular mechanism by which this network operates is unknown.

**A potential link between DNA replication and translation in Archaea**

Although no clear-cut experimental data suggest the existence of a mechanism that could link protein synthesis to DNA replication in Archaea, we recently obtained indirect evidence for such a mechanism by analyzing the environment of genes encoding DNA replication proteins in archaean genomes [2]. We observed a large cluster of seven consecutive genes encoding both DNA replication and translation proteins, which is conserved in several genomes of Crenarchaeota and partly conserved in most other archaean genomes [2]. This cluster (dubbed PPsGLSIN) includes the genes encoding PCNA (the clamp that tightly tethers several DNA replication and repair proteins to DNA), PriS (the small subunit of the DNA primase), Gins15 (one of the two subunits of the archaean GINS complex, which is involved in DNA replication initiation and elongation), the ribosomal protein L44E, the ribosomal protein S27E, the alpha subunit of the translation
Figure 1. Conserved gene associations in archaeal genomes suggest a functional coupling between DNA replication and protein translation. These two panels illustrate the conserved genomic environment of three and one DNA replication genes across different archaeal genomes. (a) A set of genes encoding three DNA replication proteins (PCNA in red; the sliding clamp; PriS in dark blue; the small subunit of the DNA primase; and Gins15 in pink, a subunit of the GINS complex) are often contiguous to four genes coding for proteins implicated in translation (the ribosomal proteins L44E (in gray) and S27E (in lemon green), and the alpha subunit of the initiation factor IF-2 (in dark brown)) or in the maturation process of the ribosome (Nop10 in tan). Only the most representative genomic neighborhoods are shown, but various alternative versions of this gene cluster are present in other archaeal genomes (for details, see Ref. [2]). The gene encoding the large subunit of the DNA primase (PriL in turquoise) is sometimes observed in this gene association. An alternative version of this cluster, including the two genes encoding the archaeal DNA polymerase D (in yellow and light brown), is present in the genome of Candidatus \textit{Korarchaeum cryptofilum}. (b) Another genomic association between one or two DNA replication genes and a gene encoding a protein involved in translation has been observed in a few euryarchaeal genomes. The genes encoding MCM (in cyan) and Gins23 (in purple) are contiguous in the majority of archaeal genomes that harbour a clear gins23 homologue, except in most genomes of Thermoproteales. In a few euryarchaeal genomes, the pair of genes \textit{mcm}-\textit{gins23} or only the \textit{mcm} gene co-localize with the gene encoding the beta subunit of the translation initiation factor aIF-2 (in magenta). Each gene is designated after the name of the protein it encodes. Some genomic contexts are identical from one genome to another; in this case, the name of the species or group of species (order) harboring the same gene arrangement is listed and the number of genomes is indicated between parentheses. White arrows correspond to functionally unrelated genes. The two mesophilic archaea \textit{Cenarchaeum symbiosum} and \textit{Nitrosopumilus maritimus} are considered as members of a tentative new archaeal phylum called the \textit{Thaumarchaeota} [40]. The hyperthermophilic archaean \textit{Candidatus Korarchaeum cryptofilum} is a member of the candidate phylum \textit{Korarchaeota} [41]. C, Crenarchaeota; DP1, DNA polymerase protein 1; DP2, DNA polymerase protein 2; E, Euryarchaeota; K, Korarchaeota; MCM, minichromosome maintenance; sp., species; T, \textit{Thaumarchaeota}. Reproduced and modified with permission from Ref. [2].
initiation factor aIF-2, and Nop10 (a protein involved in ribosome biogenesis) (Figure 1a). We recently proposed that this conserved cluster could indicate the existence of unique functional connections between DNA replication and translation in Archaea [2].

It is worth noting that the gene encoding the beta subunit of aIF-2 is adjacent to the gene encoding Mcm in seven euryarchaeal genomes (Figure 1b). The presence of DNA replication genes in the vicinity of two of the three genes encoding aIF-2 is remarkable: it suggests that aIF-2 is a good candidate for a participant in this putative regulatory network that couples DNA replication and translation. Interestingly, \textit{Pyrococcus horikoshii} aIF-2 might undergo phosphorylation [21]; eIF-2 phosphorylation is a key regulatory mechanism in eukaryotic protein synthesis (reviewed in [22]). Finally, the presence in the PPsGLSIN cluster of Nop10, a key component of the rRNA maturation apparatus in both Archaea and Eukarya [23–25], points again to a connection between DNA replication and ribosome biogenesis [19].

In our opinion, this PPsGLSIN cluster has not been preserved in Archaea solely because the genes are co-regulated. Instead, we contend that this gene arrangement has been conserved during evolution because some of the encoded proteins interact with one another [26,27]. As the association of the genes encoding L44E, S27E and Gins15 is the most highly conserved, we speculate that the connection is mediated by these proteins. As Gins15 is part of the GINS complex, which in turn associates with the MCM complex during DNA replication [28,29], we can imagine

**Figure 2.** Putative model for a dynamic coupling of DNA replication and translation processes. This cartoon illustrates a model for a functional coupling between DNA replication and translation based on the dynamic association of one or two ribosomal proteins with one key component of the DNA replication apparatus. During normal growth conditions, L44E (in grey) and S27E (in lemon green) participate in translation as components of the ribosome machinery (black arrow). Meanwhile, Cdc6 (in yellow), Gins15 (in pink) and Gins23 (in purple) as part of the GINS complex, and MCM (in cyan) are involved in DNA replication initiation (black arrow). Our model posits that one or two ribosomal proteins (L44E and S27E) display extraribosomal functions that interfere with DNA replication when the cell encounters stress conditions (e.g. amino acid starvation) or when the cell reaches the stationary phase. If these specific conditions arise (marked by the lightning symbol), the translation process halts and the ribosomal proteins S27E and/or L44E dissociate from the ribosome and become free to interact with Gins15, a core component of the replisome (blue lines). This association would lead to a hijacking of the GINS complex and its main molecular partner at the DNA replication fork, MCM, the replicative helicase. Therefore, the formation of this protein complex would inhibit the initiation and/or the elongation of DNA replication, as the participation of MCM and GINS in DNA replication is compromised in response to their association with L44E and/or S27E (red line). Protein synthesis would not be able to resume as long as L44E and/or S27E participate in this complex (red line). Eventually, the formation of this complex containing both ribosomal proteins and DNA replication proteins would lead to the cellular relocalization of MCM to the ribosomal operon region (rDNA operon) [31] (black arrow with a question mark). MCM binding to this region would result in transcription downregulation (symbolized by the magenta cross) of the rDNA operon. When conditions are favorable again, the two ribosomal proteins reintegrate into the ribosome and translation resumes. Accordingly, MCM and GINS again become free to assemble at the DNA replication fork. The representation of the ribosomal operon is indicative, and is based on the arrangement of genes in the \textit{Pyrococcus abyssi} genome.
suggesting that RPS27L participates in the p53-dependent DNA-damaging agent VP16 (etoposide), RPS27L is only three amino acid differences from RPS27), whose human homolog of S27E dubbed RPS27L (for RPS27-like; zinc-finger phosphoprotein that binds duplex DNA replication proteins across the genome of the archaeon P. abyssi via chromatin immunoprecipitation on chip [31]. We noticed that MCM preferentially binds the replication origin in the exponential growth phase, but shifts to the ribosomal operon in the stationary phase [31]. The consequences of this shift remain elusive, but it is tempting to speculate that MCM binding to the rDNA operon inhibits ribosome biogenesis in stationary phase, preventing its participation in replication initiation. Moreover, this experimental observation points to a link between the MCM complex and ribosome biogenesis in Archaea [31], which is reminiscent of the physical association between Mcm6p and the Obg-containing complex in Eukarya [19].

**Is this putative functional coupling linked to cancer formation in Eukarya?**

Remarkably, all the genes belonging to PPsGLSIN cluster, as well as the genes encoding Gins23 and MCM, have eukaryotic homologs, and all but one are absent from bacteria (PCNA is a distant relative of the beta subunit, the bacterial sliding clamp). Therefore, we speculate that the underlying functional connection between DNA replication and translation might be conserved from archaea to human. Interestingly, recent observations concerning the eukaryotic homologs of the two archaeal ribosomal proteins S27E and L44E provide support for this model.

The human homologs of S27E (known as RPS27 or MPS-1—from metallopanstimulin 1—; called RPS27 here) and L44E (RPL36A) are both involved in the control of cell growth and are linked to cancer formation [32,33]. However, the mechanism of action of these two ribosomal proteins in tumorigenesis remains elusive [32,33]. RPS27 was characterized biochemically as a nuclear zinc-finger phosphoprotein that binds duplex DNA in vitro [34]. It is therefore tempting to speculate that RPS27 binds the replication fork and/or regulatory sequences involved in DNA replication initiation. Interestingly, a second human homolog of S27E dubbed RPS27L (for RPS27-like; only three amino acid differences from RPS27), whose expression is induced by p53, was discovered recently by two independent groups [35,36]. Following treatment with the DNA-damaging agent VP16 (etoposide), RPS27L is recruited to a subset of DNA breaks where it forms foci, suggesting that RPS27L participates in the p53-dependent DNA damage response [36]. Depletion of RPS27L in p53 wild type cells results in increased cell death, which correlates with aberrant DNA damage checkpoint control. Indeed, whereas DNA synthesis decreases abruptly in normal cells following treatment by the DNA-damaging agent adriamycin, cells depleted for RPS27L continue to replicate their genome [36]. Taken collectively, these data suggest that RPS27L prevents replication forks from moving through damaged DNA, thereby preserving genome integrity.

Remarkably, all the data described above are compatible with the idea that the eukaryotic homologs of archaeal ribosomal proteins L44E and S27E are involved in the coupling between protein synthesis and DNA replication that is postulated to exist in Archaea (Figure 2). Moreover, we believe that if such an interplay does exist in Eukarya, it is likely that its deregulation would result in unbalanced cellular activity that would eventually lead to unrestricted proliferation or untimely programmed cell death. Thus, it will be important to test the effect of L44E and S27E (and their human counterparts RPL36A and RPS27) on archaeal (and eukaryotic) DNA replication proteins in various model systems.

**Concluding remarks**

Duplication of genetic material (DNA replication) and synthesis of proteins (translation) are central to cell perpetuation and therefore must be tightly regulated. Yet, the potential molecular interfaces between these two fundamental processes remain poorly investigated. We propose that a coupling between DNA replication and translation might have emerged early in cellular evolution. It is worth noting that indications for a tight coupling of DNA replication and protein synthesis were obtained in parallel by researchers working on model organisms from the three domains of life. Such findings reinforce the idea that there is much more to be learned from comparative biochemistry. Of course, the precise mechanisms of the coupling are likely to be different in each domain (e.g. most proteins of the eukaryotic Yph1 complex [19] lack archaeal homologs and the stringent response is specific for bacteria), but common themes or even common mechanisms could exist. The study of the Obg GTPases in Archaea, Bacteria, and Eukarya could be especially rewarding because these proteins are highly conserved [17]. For instance, it will be interesting to determine if Nog1p and its homologs control DNA replication by sensing the nucleotide pool in Archaea and Eukarya, as hypothesized for Obg in Bacteria; indeed such a mechanism has been documented in yeast, although the protein(s) involved is unknown [37]. The bacterial-type stringent response is absent from Archaea and most Eukarya. Consequently, the long-standing search for (p)ppGpp in eukaryotes has proven unsuccessful (reviewed in [38]), and no RelA-SpoT homolog has been identified in Archaea or Eukarya (with the notable exception of plants [10]). Yet several lines of evidence suggest that mechanisms analogous to the bacterial stringent response exist in these two domains of life. In particular, pseudomonic acid (an antibiotic that prevents tRNA charging, thereby mimicking the effect of amino acid starvation) inhibits...
rRNA synthesis in the archaeon *Sulfolobus acidocaldarius* [39]. It will be exciting to identify alarmones and proteins involved in this archaeal 'stringent response' and to determine if their induction also inhibits DNA replication. The similarity between archaeal and eukaryotic translation and DNA replication systems suggests that archaeal species could be excellent models to identify proteins that might operate in a eukaryotic stringent response and to investigate possible connections between this response and DNA replication. The eukaryotic Obg-like proteins [19], as well as the proteins encoded by the PPsSLSIN cluster uncovered by comparative genomics [2], are good candidates for participants in such critical cellular regulatory networks. The realization that a strong connection is likely to exist between DNA replication and translation is timely at the dawn of the system biology era, as biologists now focus on the integration of molecular systems at the cellular level.

**Acknowledgements**

Our work is supported by grants from the Japan Society for Promotion of Sciences to J.B., from the Canon Foundation to R.F., and from the Association pour la Recherche contre le Cancer to P.F.

**References**