Translation initiation of the replication initiator *repB* gene of promiscuous plasmid pMV158 is led by intrinsic non-SD sequences

Celeste López-Aguilar[1], José A. Ruiz-Masó[1], Tania Samir Rubio-Lepe, Marta Sanz and Gloria del Solar*.

Molecular Microbiology and Infection Biology Department. Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas. Ramiro de Maeztu 9, E28040-Madrid, Spain.

*Corresponding author. Address: Centro de Investigaciones Biológicas, Ramiro de Maeztu 9, E28040-Madrid, Spain. Phone: (+34) 918373112. Fax: (+34) 91 5360432. *E-mail address*: gdelsolar@cib.csic.es (G. del Solar).

[1] These authors contributed equally to this work.

Keywords: promiscuous plasmid pMV158, *copG-repB* mRNA, translational coupling, *repB* translation efficiency, ARBS, translation initiation region.

Abstract

RepB is the pMV158-encoded protein that initiates rolling-circle replication of this promiscuous plasmid. Availability of RepB is rate-limiting for the plasmid replication process, and therefore the *repB* gene encoding the protein is subjected to strict control. Two trans-acting plasmid elements, CopG and the antisense RNAII, are involved in controlling the synthesis of the initiator at the transcriptional and translational level, respectively. In addition to this dual control of *repB* expression that senses and corrects fluctuations in plasmid copy number, proper availability of RepB also relies on the adequate functionality of the transcription and translation initiation regulatory signals. Translation of *repB* has been postulated to depend on an atypical ribosome binding site that precedes its start codon, although such a hypothesis has never been proved. To define sequences involved in translation of *repB*, several mutations in the translation initiation region of the *repB* mRNA have been characterized by using an *Escherichia coli in vitro* expression system wherein the synthesis of RepB was detected and quantified. We showed that translation of *repB* is not coupled to that of *copG* and depends only on its own initiation signals. The atypical ribosome binding site, as it was defined, is not involved in translation initiation. However, the sequence just upstream of the *repB* start codon, encompassing the proximal box of the atypical ribosome binding site and the four bases immediately downstream of it, is indeed important for efficient translation of *repB*. The high degree of conservation of this sequence among the *rep* genes of plasmids of the same pMV158 family supports its relevancy as a translation initiation signal in mRNAs without a recognizable Shine-Dalgarno sequence.

1. Introduction

PMV158 is a small (5536 bp) multicopy promiscuous plasmid that initiates its rolling circle replication by the plasmid-encoded initiator RepB interacting specifically with the double-strand origin (*dso*) of replication. RepB introduces a site-specific cleavage on one plasmid DNA strand within the *dso* and yields a free 3'-OH end that serves to prime the synthesis of the plasmid leading strand (del Solar et al., 1993; Moscoso et al., 1995). Thus, the frequency of replication would be limited only by the availability of the Rep protein, which should be subjected to tight control to maintain a given plasmid copy number (Novick, 1987). The mechanism controlling the replication of pMV158 involves two trans-acting plasmid elements, namely an antisense (or countertranscript) RNA (RNAII) and the transcriptional repressor protein CopG, both regulating the synthesis of the initiator RepB protein (del Solar et al., 1995). The *copG* and *repB* genes constitute an operon transcribed from the single P_{cr} promoter (del Solar et al., 1990). CopG represses synthesis of the *copG-repB* mRNA by binding to its operator, which includes the P_{cr} promoter (Fig. 1), hence preventing the access of the RNA polymerase to it (Hernández-Arriaga et al., 2009). At the translational level, expression of *repB* is inhibited by direct pairing of antisense RNAII, whose synthesis is directed by promoter P_{ctII} (Fig. 1), with its complementary region in the bicistronic *copG-repB* mRNA. The combination of both regulatory elements, CopG and RNAII, produces a synergistic effect that increases their efficacy in maintaining the plasmid copy number within a narrow range.

Sequence analysis of the *copG-repB* operon revealed only a typical Shine-Dalgarno (SD) sequence for the *copG* gene. The small intergenic region, which encodes the antisense RNA (Fig. 1), was proposed to contain what was termed an atypical ribosome binding site (ARBS; (Lacks et al., 1986) for *repB* translation. Designation of ARBS was based on the presence of a conserved sequence 9 or 10 nts upstream from the start codon of *repB* and of the *dpnM* and *dpnA* methylase genes from pneumococcus, with no apparent complementarity with either 16 S or 23 S rRNA (de la Campa et al., 1987; Mannarelli et al., 1985). The ARBS consists of consensus upstream (5'-AATTTCT-3') and downstream (5'-TATA-3') regions 4 or 5 nts apart from each

other, which were designated as distal and proximal boxes, respectively, because of their distance to the start codon (Fig. 1). Comparison of the sequence upstream from the *rep* genes of various plasmids of the pMV158 family with that of *repB*, led the authors to propose that translation of the initiator genes also depended on ARBS-type signals (Kim et al., 2008). However, the role of the ARBS in initiation of translation has never been proved for any gene. Interestingly, in the case of the *dpnM* gene, the proposed ARBS sequence was in fact a -10 extended promoter, with no required -35 site, from which the transcription was initiated at the methylase start codon, thereby producing an mRNA without any apparent ribosome-binding site (Sabelnikov et al., 1995). This does not seem to be the case with the ARBS located upstream of the *repB* gene, since transcription from an internal promoter located in the *copG-repB* intergenic region was not observed by S1 nuclease protection or *in vitro* transcription experiments (del Solar et al., 1990). In the case of *repB*, substitution of the four bases 3'-adjacent to the ARBS proximal box (mutation m in the hereafter termed 3'-ar region), was reported to affect *in vitro* translation of the gene under conditions wherein RNAII was not synthesized (del Solar et al., 1997). This suggested that the translation initiation signals of *repB* should be extended from the proposed ARBS proximal box towards the ATG of the gene.

In this work, we have generated various deletion mutants to determine whether expression of *repB* is translationally coupled to that of *copG*. We have found that *repB* carries its own translation initiation signals and that its expression does not depend on the translation of *copG*. In addition, we have generated several mutants by substitution of different sequences in the region upstream of the *repB* initiation codon, in order to determine the involvement of the postulated ARBS and other nearby sequences in translation initiation of the replication initiator gene. Only changes in the ARBS proximal box and in the 3'-adjacent region produced a significant decrease in the synthesis of the protein, confirming the importance of these sequences in the translation efficiency of *repB*, and calling in question the involvement of the ARBS, as it was defined, in initiation of translation.

2. Materials and methods

2.1. Bacterial strain, plasmids and oligonucleotides

E. coli DH5α™-T1^R strain was the host for site directed mutagenesis. Construction of plasmid pALT7:crcat is described in (del Solar et al., 1997; Gomis-Rüth et al., 1998). The oligonucleotides used for mutagenesis are listed in Table 1.

2.2. Construction of deletion mutants and site directed mutagenesis

Deletion mutants were created by following an inverse PCR procedure. Two sets of divergent primers were used to generate the mutants ΔA4-m (ΔA4-F, ΔA4-SD-R) and ΔA4^{SD}-m (ΔA4-F, ΔA4-R), employing pALT7:crcat as template. After amplification with the Phusion polymerase (Finnzymes), the resultant DNA fragments were gel-purified and incubated with T4 DNA ligase to generate circular molecules that were used to transform competent *E. coli* cells. To allow ligation-sealing of the cyclized DNA molecules, the primers were phosphorylated at their 5' end before being used in the amplification reaction. The short deletion in mutant ΔA1-m was introduced by site directed mutagenesis as described below.

The GeneTailor™ System (Invitrogen) was used to perform site directed mutagenesis to generate the ΔA1-m, loop-m, AU_1-m and AU_2-m, D-m, P-m and m mutants. The plasmid pALT7:crcat was used as DNA template and the overlapping primers (Table 1) were designed following the manufacture's specifications. The expected mutations were confirmed by DNA sequencing and the resulting mutant plasmids were purified with a Midi-prep reaction kit (Genomed).

2.3. In vitro synthesis of mRNA and proteins

T7 RNA polymerase-directed mRNA synthesis was performed with template mutant plasmid DNAs digested at the single NcoI site located within the *cat* gene (Fig. 2). Digested DNAs were treated with proteinase K and SDS

as described in (del Solar et al., 1997) and used as templates for transcription. The transcription assays were performed using T7 RNA polymerase from Roche, according to the manufacturer's protocol. The transcription products were visualized and quantified, with the aid of the Quantity One software (Bio-Rad), from ethidium bromide-stained agarose gels wherein known amounts of a control RNA were loaded.

For translation assays, 2 pmol of *in vitro* synthesized RNAs were added to S-30 *E.coli* cell-free extracts especially designed for linear templates (Promega). To prevent synthesis of antisense RNAII encoded by the template plasmid DNA, translation was uncoupled from transcription by incubating with rifampicin (25 μg ml^{-1}). *De novo*-synthesized proteins were labeled with [^{35}S]-methionine, analyzed by SDS-PAGE, and the radioactivity in the bands was detected and quantified with the aid of a FLA-3000 (FUJIFILM) imaging system and the Quantity One software (Bio-Rad). Efficiency of *repB* translation was calculated from at least three independent experiments, employing two different stocks of *in vitro* synthesized RNAs. Genes *copG* and *cat* were used as internal controls to correct for differences in the translation efficiency of different RNA samples. Whenever possible, values of radioactivity in RepB were normalized with respect to those in CopG, since this protein was a major product in the *in vitro* translation assays, which facilitated its quantification. Similar results were obtained, notwithstanding, when RepB values were normalized with respect to the levels of truncated (T) Cat (chloramphenicol acetyltransferase).

2.4. Prediction of RNA secondary structures

RNA structures were predicted by using the Mfold Web Server (http://mfold.rna.albany.edu, Zuker, M. 2003).

3. Results and Discussion

3.1. Gene repB *is not translationally coupled to* copG

As mentioned above, *copG* and *repB* genes are co-transcribed to form a bicistronic *copG-repB* mRNA wherein the two open reading frames (ORFs) are

separated by only sixty bases. Both the proximity between the two ORFs and the absence of a typical SD sequence just upstream of *repB* prompted us to speculate that the translation of this essential gene might depend on the translation of the previous *copG* gene. Thus, a translational coupling mechanism can be conjectured by which the ribosome, after terminating translation of the upstream gene, would search for the *repB* translation initiation signal and restart translation. The existence of such a translational coupling was investigated by analyzing the translation efficiency of several deletion mutants. We had previously constructed plasmid pLS1ΔA4, a pMV158 deleted derivative that lacks the entire *copG* gene but conserves the *copG* SD sequence and the *repB*-ARBS (Fig. 1) (del Solar and Espinosa, 1992). This construction was successfully established in *Streptococcus pneumoniae*, demonstrating that *repB* is translated even though the previous gene *copG* is not. However, it remained unclear whether translation of *repB* in pLS1ΔA4 was initiated from its own translation initiation signals or from the *copG* SD sequence, which, as a consequence of the deletion, is closer to *repB*.

Now, we have constructed various *copG* deletion mutants of plasmid pALT7:crcat (del Solar et al., 1997) aimed at analyzing *repB* translation from *in vitro* synthesized transcripts. Plasmid pALT7:crcat contains a promoterless *copG-repB* operon and a *cat* gene placed under the control of the Φ*10* promoter recognized by the T7 RNA polymerase (see Fig. 2). After digestion with NcoI, the linearized plasmid was used as a template for the synthesis of a *copG-repB*-truncated *cat* mRNA that, in turn, was employed for *in vitro* translation. It is worth noting that the *E. coli* cell-free extracts used for the translation assays were treated with rifampicin to uncouple translation from transcription. This avoided transcription from the P_{ctII} promoter and, therefore, the synthesis of antisense RNAII, which is known to inhibit *repB* translation. Three different deletion mutants were assayed in this system: ΔA4-m, which shows the same deletion as plasmid pLS1ΔA4; $\Delta A4^{SD}$-m, whose deletion also includes the *copG* SD sequence; and ΔA1-m, which lacks the first three codons of *copG* (Fig. 2). An expected consequence of all three deletions was the absence of expression of *copG.* In spite of this, the yield of RepB was similar (ΔA1-m) or even higher (ΔA4-m and $\Delta A4^{SD}$-m) than that obtained from the wild type construction (Fig. 3A). This demonstrates that *repB* translation can occur in the absence of *copG*

(ΔA4-m and ΔA4^{SD}-m mutants) or in the presence of an untranslatable *copG* gene (mutant ΔA1-m). Interestingly, the largest increase in the efficiency of RepB synthesis was observed in the mutant that lacks the entire *copG* gene along with its SD (ΔA4^{SD}-m), thus showing that translation of *repB* in ΔA4-m does not require nor is improved by the neared SD of *copG*. At present, we do not have any convincing explanation for the moderate (50%) and large (200%) increases in the *repB* translation efficiency of the ΔA4-m and ΔA4^{SD}-m mutants, respectively. It is worth noting, however, that the 5'-end of the *copG-repB-cat* mRNA is located much closer to the *repB* translation initiation regions in these two mutants than in the wild type (Fig. 2). This might facilitate searching of the *repB* translation initiation signals by the 30S ribosomal small subunit after its initial unspecific binding.

Together, all the above results allow ruling out the existence of translational coupling between the two genes of the *copG-repB* operon, and confirms the functionality of an intrinsic translation initiation region of *repB*.

3.2. Identification of important sequences for translation of the repB *gene*

Plasmid pMV158 was originally isolated from *Streptococcus agalactiae* and due to its promiscuity has been successfully transferred to several Gram-positive bacteria from Firmicutes and Actinobacteria. Even in the absence of a typical SD sequence, the *repB* translation-initiation signals are recognized in the Gram-negative bacterium *E. coli* as shown (i) by the ability of pMV158 to replicate *in vivo* and *in vitro* using the machinery provided by this bacterium, and (ii) by the functionality of the purified RepB protein, which is overproduced in an *E. coli*-based expression system that uses the original translation-initiation signals of *repB*. In fact, determination of the amino-terminal end of the purified protein showed that the start codon used in *E. coli* is also the one used in *S. pneumoniae* (de la Campa et al., 1990; Ruiz-Masó et al., 2004). These results reinforce the suitability of the *E. coli* based *in vitro* expression system used in this work to analyze the expression of *repB* in several translation initiation region mutants.

Using computer simulations with the Mfold algorithm (Zuker, 2003), the region of the *copG-repB* mRNA around the initiation codon of *repB* was

predicted to form three stem-loop structures (hairpins 1, 2, and 3 in Fig. 4A). The initiation codon of *repB* is located in the single-stranded portion between the stem-loop 2 and the base of the stem-loop 3, whereas the ARBS proximal box forms part of the stem-loop 2. A G-tract was identified in the lower stem portion of the stem-loop 2, just upstream from the proximal box of the ARBS. The single-stranded portion between the hairpins 1 and 2 constitutes an AU-rich region that includes the distal box of the ARBS. A series of mutations were introduced in the expression vector pALT7:crcat to analyze the involvement of some of these sequences in the translation efficiency of *repB*. The entire sequence of the distal and proximal boxes of the ARBS was changed by a random sequence in the mutants D-m and P-m, respectively (Fig. 2). Previously, the 3'-ar region, which is 3'-adjacent to the ARBS proximal box, had been substituted to generate the m mutation (del Solar et al., 1997). Furthermore, the stretch of nine bases just downstream of the stem-loop 1 was entirely changed by GC-rich sequences in the mutants AU_1-m and AU_2-m, while the loop sequence of the stem-loop 1 was randomly changed to generate the mutant loop-m. The sequences to be changed were chosen because of their proposed involvement in *repB* translation (D-m, P-m and m mutants), or because they were complementary to the RNAII single-stranded regions that participate in the initial recognition of the *copG-repB* intergenic region of the target mRNA (AU_1-m, AU_2-m and loop-m mutants; unpublished results). Therefore, mutants AU_1-m, AU_2-m and loop-m were useful to determine the involvement of these sequences in the translation efficiency of *repB*.

Experiments of *in vitro* translation of the *copG-repB-cat* transcripts obtained from these mutants showed that, whereas the efficiency of RepB synthesis in the mutant D-m was similar to that in the wild type, the mutations P-m and m caused a 3-4-fold reduction in the translation of *repB* (Fig. 3B). This suggests that only the proximal box of the putative ARBS is involved in ribosome binding, and confirms previous results showing the involvement of the downstream 3'-ar in initiation of *repB* translation (del Solar et al., 1997). Mutations AU_2-m and loop-m did not alter significantly the efficiency of RepB synthesis, although a moderate but significant decrease (40%) was observed in the mutant AU_1-m (Fig. 3B). Interestingly, mutations AU_1-m and AU_2-m, despite both of them changing the upstream AU-rich region to a GC-rich sequence,

affected distinctly not only the yield of RepB, but also that of CopG (Fig. 3B). In fact, mutant AU_2-m consistently yielded a lower amount of CopG than the wild type and the other mutants. We speculate that this may result from AU_2-m-derived mRNA being prone to endonuclease cleavage at the mutated region and subsequent 3'-5' exonucleolytic degradation. Since CopG synthesis seemed to be affected, the RepB yield in mutant AU_2-m was normalized with respect to that of T-Cat.

A comparative analysis between the predicted secondary structure of the 5' leader region of the *rep* mRNA of plasmid ColE2, which also lacks a typical SD sequence, and that of *repB* revealed a similar organization of sequence and structural elements potentially involved in translation initiation of the essential genes (Fig. 4A). It has been shown that the stem-loop 1 and the AU-rich sequence downstream of it are involved in translation initiation of the *rep* gene of ColE2 (Nagase et al., 2007). It has also been suggested that these two regions could constitute a recognition site for the ribosomal protein S1, which enhances the efficiency of translation of mRNAs carrying a weak SD sequence. Furthermore, a GA cluster, which has been speculated to function as a weak SD sequence (Yasueda et al., 1994), along with the sequence that separates it from the initiation codon are important for the efficient translation of the ColE2 *rep* gene (Nagase et al., 2007). Similarly, in the case of the *repB* translation initiation region, the stem-loop 1 and the contiguous AU-rich region may promote binding of the S1 protein and, therefore, serve as a ribosome recognition site. However, neither of these elements seems to be essential for efficient translation of *repB*. In fact, the stem-loop 1 can be removed (ΔA4-m and $\Delta A4^{SD}$-m) without impairing translation of the essential gene, while the upstream (AU_1-m and AU_2-m) or the downstream (D-m) parts of the AU-rich region can be substituted by GC-rich sequences with little or no detriment to the efficiency of *repB* translation. Since stem-loop 1 along with the AU-rich region comprise the entire sequence complementary to RNAII, it follows that binding of the inhibitory RNA does not block directly any sequence element required for efficient *repB* translation, although formation of the RNA-RNA duplex might hinder the binding of the ribosome to the downstream region.

A 3-nt G-tract is observed in the *copG-repB* intergenic region of the mRNA, between the distal and proximal boxes of the ARBS (Fig. 4A). This

small G-tract shares some common features with the functionally relevant GA cluster of ColE2, namely a similar spacing respect to the start codon (13 nts), and its location in a secondary structure of the mRNA. Therefore, the G-tract is a potential, although yet untested, key element for *repB* translation. Interestingly, stem-loop 2 of the *copG-repB* mRNA, which includes the untested G-tract along with the relevant ARBS proximal box and 3'-ar, is predicted to be quite unstable (ΔG= -1.0; Fig. 4B), which may facilitate the accessibility of these important elements to the 30S ribosome subunit. Our results indicate that the efficiency of translation is most impaired by sequence substitutions of the ARBS proximal box and of the 3'-ar. Both mutations also cause the rearrangement of the predicted structure of the spacer between stem-loops 1 and 3, placing the G-tract within hairpins that are more stable than stem-loop 2 (Fig. 4B). Seizing of the G-tract in a more stable structure is also observed in mutant AU_2-m (Fig. 4B), although impairment of translation was not observed in this case. Therefore, sequence changes in the proximal region and in the 3'-ar, rather than changes in the local structure or in the accessibility of the G-tract, most likely account for the decrease in translation efficiency of P-m and m mutants.

3.3. Comparison of the translation initiation region of the rep *genes of plasmids of the pMV158 family.*

Plasmids of the pMV158 family have been shown colonizing Gram-positive bacteria from Firmicutes but also some bacteria from different phyla, like *Mycoplasma mycoides* (Mollicutes) and the Gram-negative *Helicobacter pylori* (Proteobacteria). For that reason, the alignment shown in Fig. 5 only displays sequences of the 5' regions of the *rep* genes of plasmids isolated from Firmicutes that, in principle, should share more homology in their translation initiation signals. A first inspection of the alignment revealed that none of these sequences displays a typical SD sequence. By contrast, the so-called 3'-ar region is conserved in more than 80% of the sequences, although the degree of conservation decreases in the ARBS proximal box, where only the first and the last position of the consensus sequence 5'-TATA-3' are conserved. The homology found at this region is due, at least in part, to the overlapping -10 element of the P_{ctII} promoter that is located in the opposite strand and directs

synthesis of the inhibitory antisense RNA. Interestingly, the G-tract found in pMV158 is partially conserved in the rest of the sequences, wherein the consensus is 5'-AGG-3'. Lastly, the degree of conservation of the distal box of the ARBS is much lower than the observed in the other sequence elements. This finding fits well with the results shown in this work and highlights the relevancy of the sequence immediately upstream of the start codon in the translation initiation region when a typical SD sequence is not present.

4. Conclusions

Expression of the essential *repB* gene is strictly regulated at the transcriptional and translational level. Here we present new data about the identification and characterization of the initiation signals that regulate the translation of *repB*. We show that translation of *repB* relies on its own initiation signals and, therefore, discard a possible mechanism of translational coupling to the upstream *copG* gene. Changes in the sequence of the ARBS proximal box and of the 3'-ar, but not in the ARBS distal box, produced the largest reduction in *in vitro* synthesis of RepB, demonstrating the importance of the region immediately upstream of the *repB* start codon in the efficiency of translation of *repB*, and calling into question the functionality of the postulated ARBS. No further evidences about the involvement in the translation efficiency of *repB* of other sequence and structural elements, like the stem-loop 1 or the AU-rich region, were found. Finally, the conclusions obtained from this study could be applicable to the majority of the plasmids of the pMV158 family due to the high degree of identity found at the *rep* translation initiation regions.

Acknowledgments

This study was supported by the Spanish Ministerio de Ciencia e Innovación (Grants CSD2008-00013, INTERMODS, and BFU2010-19597, PNEUMOTALK to GdS). T. S. R-L. was a recipient of a PreDoc JAE fellowship from the Consejo Superior de Investigaciones Científicas.

References

Crooks, G. E., Hon, G., Chandonia, J.-M., Brenner, S. E., 2004. WebLogo: A Sequence Logo Generator. Genome Res. 14**,** 1188-1190.

de la Campa, A. G., del Solar, G., Espinosa, M., 1990. Initiation of replication of plasmid pLS1. The initiator protein RepB acts on two distant regions. J. Mol. Biol. 213**,** 247-262.

de la Campa, A. G., Kale, P., Springhorn, S. S., and Lacks, S. A., 1987. Proteins Encoded by the *Dpn*II Restriction Gene Cassette. Two Methylases and an Endonuclease. J. Mol. Biol. 196**,** 457-469.

del Solar, G., Acebo, P., Espinosa, M., 1995. Replication control of plasmid pLS1: efficient regulation of plasmid copy number is exerted by the combined action of two plasmid components, CopG and RNA II. Mol. Microbiol. 18**,** 913-924.

del Solar, G., Acebo, P., Espinosa, M., 1997. Replication control of plasmid pLS1: the antisense RNA II and the compact rnaII region are involved in translational regulation of the initiator RepB synthesis. Mol. Microbiol. 23**,** 95-108.

del Solar, G., Espinosa, M., 1992. The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products : the antisense RNA II and the repressor protein, RepA. Mol. Microbiol. 6**,** 83-94.

del Solar, G., Moscoso, M., Espinosa, M., 1993. Rolling circle-replicating plasmids from gram-positive and -negative bacteria: a wall falls. Mol. Micobiol. 8**,** 789-796.

del Solar, G. H., Pérez-Martín, J., Espinosa, M., 1990. Plasmid pLS1-encoded RepA protein regulates transcription from repAB promoter by binding to a DNA sequence containing a 13-base pair symmetric element. J. Biol. Chem. 265**,** 12569-75.

Gomis-Rüth, F. X., Solà, M. a., Pérez-Luque, R., Acebo, P., Alda, M. T., González, A., Espinosa, M., del Solar, G., Coll, M., 1998. Overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the pMV158-encoded plasmid transcriptional repressor protein CopG. FEBS Lett. 425**,** 161-165.

441 Hernández-Arriaga, A. M., Rubio-Lepe, T. S., Espinosa, M., del Solar, G., 2009.
442 Repressor CopG prevents access of RNA polymerase to promoter and
443 actively dissociates open complexes. Nucl. Acids Res. 37**,** 4799-4811.
444 Kim, S. W., Jeong, I. S., Jeong, E. J., Tak, J. I., Lee, J. H., Eo, S. K., Kang, H.
445 Y., and Bahk, J. D., 2008. The terminal and internal hairpin loops of the
446 ctRNA of plasmid pJB01 play critical roles in regulating copy number.
447 Mol. Cells. 26**,** 26-33.
448 Lacks, S. A., López, P., Greenberg, B., Espinosa, M., 1986. Identification and
449 analysis of genes for tetracycline resistance and replication functions in
450 the broad-host-range plasmid pLS1. J. Mol. Biol. 192**,** 753-765.
451 Mannarelli, B. M., Balganesh, T. S., Greenberg, B., Springhorn, S. S., Lacks, S.
452 A., 1985. Nucleotide sequence of the *Dpn* II DNA methylase gene of
453 *Streptococcus pneumoniae* and its relationship to the *dam* gene of
454 *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 82**,** 4468-4472.
455 Moscoso, M., del Solar, G., Espinosa, M., 1995. Specific nicking-closing activity
456 of the initiator of replication protein RepB of plasmid pMV158 on
457 supercoiled or single-stranded DNA. J. Biol. Chem. 270**,** 3772-3779.
458 Nagase, T., Nishio, S.-y., Itoh, T., 2007. Importance of the leader region of
459 mRNA for translation initiation of ColE2 Rep protein. Plasmid. 58**,** 249-
460 260.
461 Novick, R. P., 1987. Plasmid incompatibility. Microbiological Reviews. 51**,** 381-
462 395.
463 Ruiz-Masó, J. A., López-Zumel, C., Menéndez, M., Espinosa, M., del Solar, G.,
464 2004. Structural features of the initiator of replication protein RepB
465 encoded by the promiscuous plasmid pMV158. Biochim. Biophys. Acta -
466 Proteins & Proteomics. 1696**,** 113-119.
467 Sabelnikov, A. G., Greenberg, B., Lacks, S. A., 1995. An extended -10 promoter
468 alone directs transcription of the *Dpn*II operon of *Streptococcus*
469 *pneumoniae*. J. Mol. Biol. 250**,** 144-155.
470 Yasueda, H., Takechi, S., Sugiyama, T., Itoh, T., 1994. Control of ColE2
471 plasmid replication: negative regulation of the expression of the plasmid-
472 specified initiator protein, Rep, at a posttranscriptional step. Mol. Gen.
473 Genet. 244**,** 41-48.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucl. Acids Res. 31**,** 3406-3415.

Figure Legends

Figure 1. Replication control region of pMV158. The nucleotide sequence between coordinates 592 and 888, encompassing the control elements *copG* and *rnaII* and the first codons of the replication initiator *repB* gene, is depicted. The relevant features of this region, namely the *copG* initiation and termination codons, the *repB* initiation codon, the -35 and -10 regions of promoters P_{cr} and P_{ctII}, and the putative ribosome binding sites previously proposed for *copG* (SD) and *repB* (ARBS), are underlined. The coding sequence of *rnaII* is shown in italics. The initiation site of *rnaII* is indicated by an arrow pointing in the direction of the transcription. The palindrome corresponding to the transcription terminator (T_{II}) of *rnaII* is indicated by opposite arrows below the sequence. Plasmid sequence deletion in pLS1ΔA4 is also shown.

Figure 2. Expression vector designed for *in vitro* transcription and translation assays. Plasmid pALTER-1 (Promega) was used to clone the genes *copG, repB* and *cat* from pJS3. This cloning places these genes under the control of the ϕ*10* promoter (box), and inactivates the P_{cr} promoter of *copG* and *repB* (del Solar et al., 1997; Gomis-Rüth et al., 1998). The resultant recombinant plasmid, pALT7:crcat, was used to express *in vitro* the genes cloned under the ϕ*10* promoter. The nucleotide sequence of the region encompassing from the ϕ*10* promoter to the beginning of *repB* is shown. Several deletions (ΔA4-m, $\Delta A4^{SD}$-m and ΔA1-m) and base changes (loop-m, AU_1-m, AU_2-m, D-m, P-m and m) where introduced in pALT7:crcat to study their influence in the expression of *repB*. The coordinates of the sequence corresponding to pMV158 are indicated.

Figure 3. Effect of the mutations in the translation efficiency of *repB*. Recombinant vectors with the mutations ΔA4-m, $\Delta A4^{SD}$-m, ΔA1-m, loop-m, AU_1-m, AU_2-m, D-m, P-m and m were used to express *in vitro* the genes cloned under the T7 RNA polymerase-dependent ϕ*10* promoter (see Fig. 2). Template

DNAs used for *in vitro* mRNA synthesis were linearized with NcoI (see Fig. 2). *In vitro* translation was performed with S30 *E. coli* cell-free extracts (Promega). *De novo*-synthesized proteins were ^{35}S-labeled, separated on 19% SDS-PAA gels, and visualized and quantified using a phosphorimager system. The positions of the protein bands corresponding to RepB, T-Cat and CopG are shown. RepB synthesis was quantified and corrected for the amount of CopG or T-Cat synthesized. The efficiency of *repB* translation in the different mutants relative to the wild type construction was determined. **(A)** The efficiency of RepB synthesis, normalized with respect to T-Cat, was 122(±11)%, 148(±9)% and 300(±7)% for the deletion mutants ΔA1-m, ΔA4-m and ΔA4^{SD}-m, respectively. **(B)** The efficiency of RepB synthesis, corrected for CopG, was close to the wild type in the mutants D-m and loop-m (93(±16)% and 93.8(±20)%, respectively) and significantly lower in the mutants P-m, m and AU_1-m (26(±11)%, 33(±8)% and 58.7(±11)%, respectively). In the case of the mutant AU_2-m, the relative translation of *repB* (109(±22)%) was obtained after correction for T-Cat due to the observed reduction in the synthesis of CopG respect to the wild type.

Figure 4. Sequence and structural elements of the *repB* translation initiation region. (A) Predicted secondary structure of the intergenic region of the *copG-repB* mRNA. The start codon of *repB* is underlined. The position where the sequence complementary to RNAII starts is indicated by an arrow. The AU-rich region and the G-tract are indicated by shadowed boxes. The sequence regions altered in the loop-m, AU_1-m, AU_2-m, D-m, P-m and m mutants are also indicated. *Inset:* schematic representation of the structure of the 5' end portion of the *rep* mRNA of ColE2 (Nagase *et al.,* 2007). Important regions for initiation of *rep* translation, like the AU-rich region, the GA cluster, the start site of the regulatory RNAI and the start codon, are indicated. For clarity, the stem-loop I was eliminated from the structure, so that the stem-loops II, III and IV in the original article correspond to the hairpins 1, 2 and 3 represented in the figure. **(B)** Predicted secondary structure of the intergenic region of various *copG-repB* mRNA mutants. A cartoon format is used to display the hairpins composing the most energetically favorable secondary structure predicted for each mutant. Thick dashes in the stem of the hairpins

indicate G-C base pairs, whereas thin dashes represent A-U and G-U pairs. Hairpins 1, 2 and 3 correspond basically to those shown in panel **A**, which have ΔG values (obtained by using the Mfold algorithm) of -12.4, -1.0 and -5.1, respectively. Hairpins 1 and 3 are conserved in all the mutants, whereas the RNA located between these two hairpins is predicted to adopt different secondary structures. The hairpin that seizes the G-tract in every mRNA variant is displayed in bold type dashes, and its ΔG value is indicated. The stem-loop located between hairpins 1 and 2 in the D-m mutant has a ΔG value of -2.6. The *repB* start codon, which is 5'-adjacent to hairpin 3, is boxed. G residues in the G-tract are displayed by small open circles.

Figure 5. Sequence alignment of the putative translation initiation region of the rep genes of the pMV158 plasmid family. The start codons for the Rep proteins are aligned at the right end of the figure. The sequence of the distal (D) and proximal (P) boxes of the postulated pMV158-*repB* ARBS, as well as the G-tract and the 3'-ar, are indicated on top of the alignment. The conserved sequences were shadowed with different intensities depending on the degree of conservation at each position; dark grey >80%, grey >60% and light grey >40%. Weblogo (Crooks et al., 2004) was used to generate a logo with the sequences included in the alignment. The height of each letter is proportional to the frequency of the corresponding nucleotide, whereas the overall height of each stack of letters is proportional to the sequence conservation, measured in bits, at that position.

Tables

Table 1. List of oligonucleotides

oligonucleotide	size (nt)	5´-3´sequence	used to construct
ΔA1-F	37	GTTTTATAAAATTTTGAGAGGTGAAGATTGACGATAA	ΔA1-m
ΔA1-R	37	TCACCTCTCAAAATTTTATAAAACTATAGTAGCATAA	
ΔA4-F	21	GGCTAAAGTCAAACATTTCTT	ΔA4-m and $\Delta A4^{SD}$-m
ΔA4-R	21	CTATAGTAGCATAACCGTGCA	
ΔA4-SD-R	27	CCTCTCAAAATTTTATAAAACTATAGT	
Dm-F	39	GGCACTGGCTAAAGTCAAACCGACTCTGGGTATATTATA	a.r.b.s distal box mutant (D-m)
Dm-R	30	GTTTGACTTTAGCCAGTGCCTGCCAGCACG	
Pm-F	37	CTAAAGTCAAACATTTCTTGGGCGACTTATACTTTAT	a.r.b.s proximal box mutant (P-m)
Pm-R	30	CCCAAGAAATGTTTGACTTTAGCCAGTGCCTG	
loop-F	37	GAAAAATAAAAAAAGCCGTGCGAAGTTGCACTGGCTA	loop mutant (loop-m)
loop-R	30	GCACGGCTTTTTTTATTTTTCTTGACCTTT	
AU_1m-F	38	CTGGCAGGCACTGGCGACGACGACACATTTCTTGGGTA	AU region mutants (AU_1-m/AU_2-m)
AU_2m-F	38	CTGGCAGGCACTGGCGCCGGGAGCACATTTCTTGGGTA	
AUm-R	25	GCCAGTGCCTGCCAGCACGGCTTTT	

*mutations are underlined

Figure 1

Click here to download high resolution image

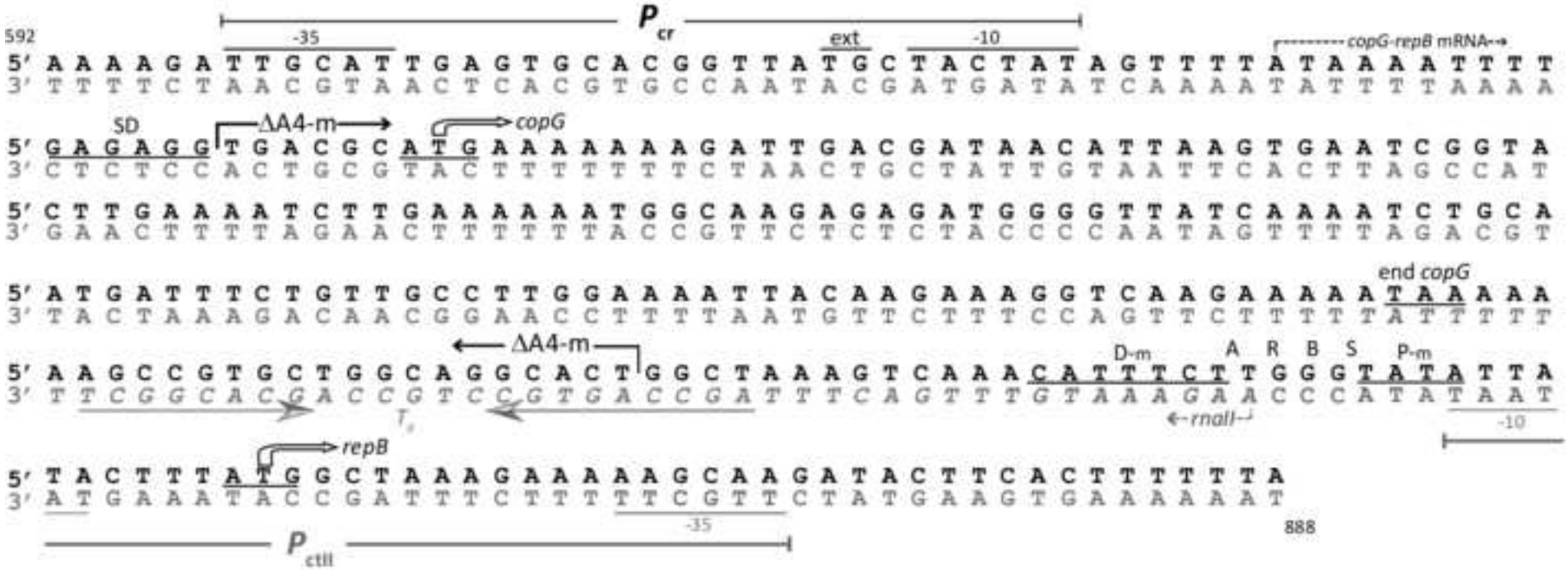

Figure 2

Click here to download high resolution image

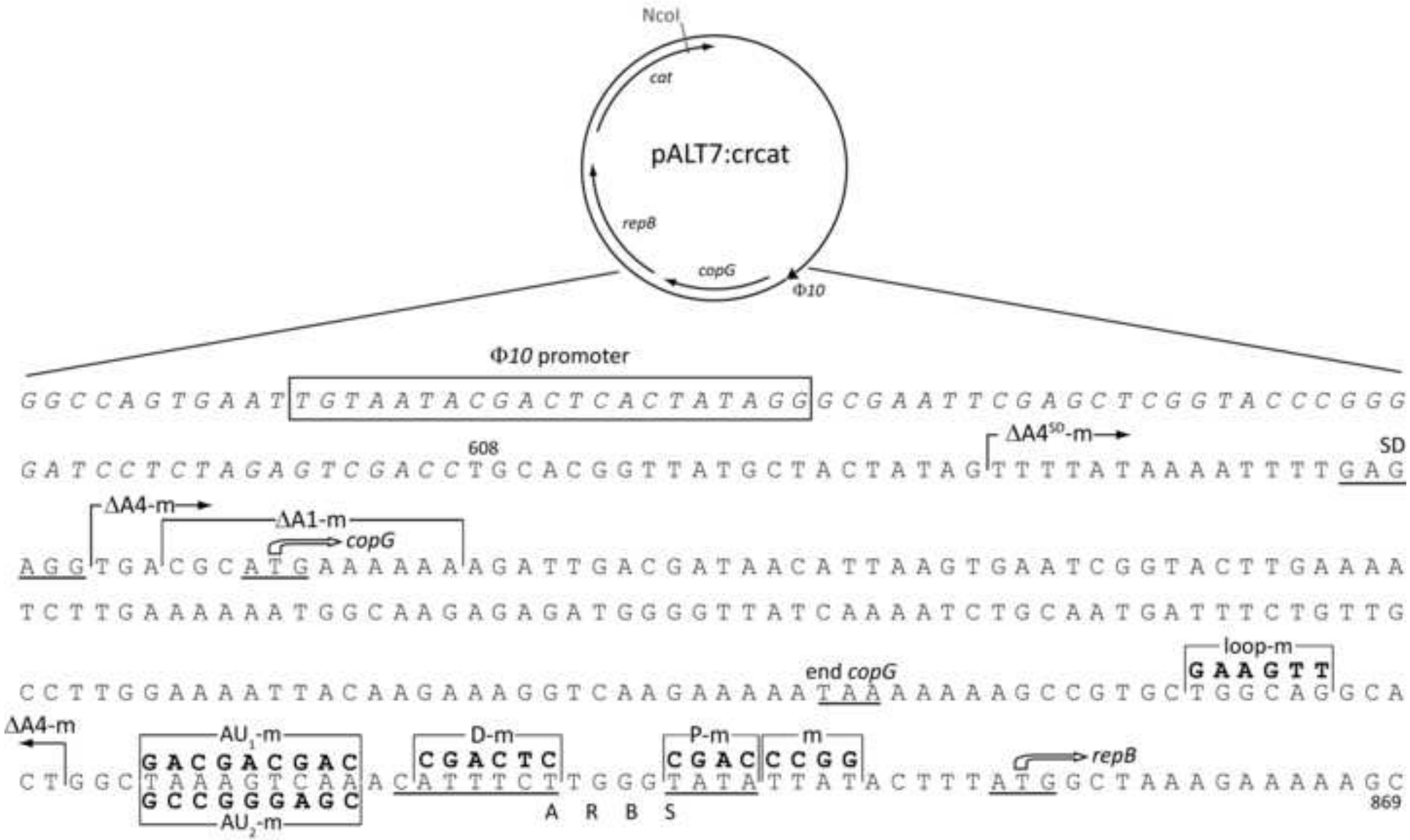

Figure 3

Click here to download high resolution image

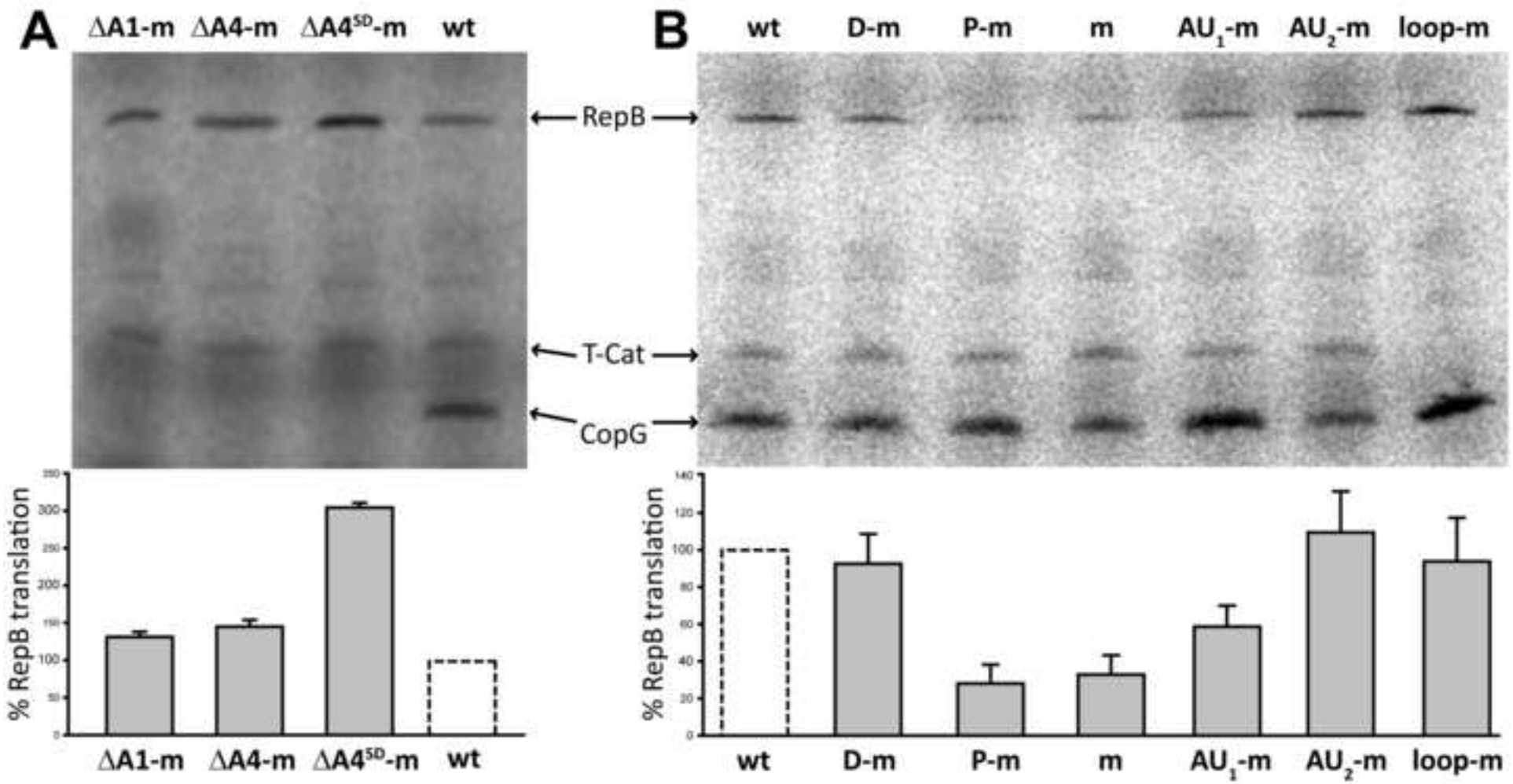

Figure 4

Click here to download high resolution image

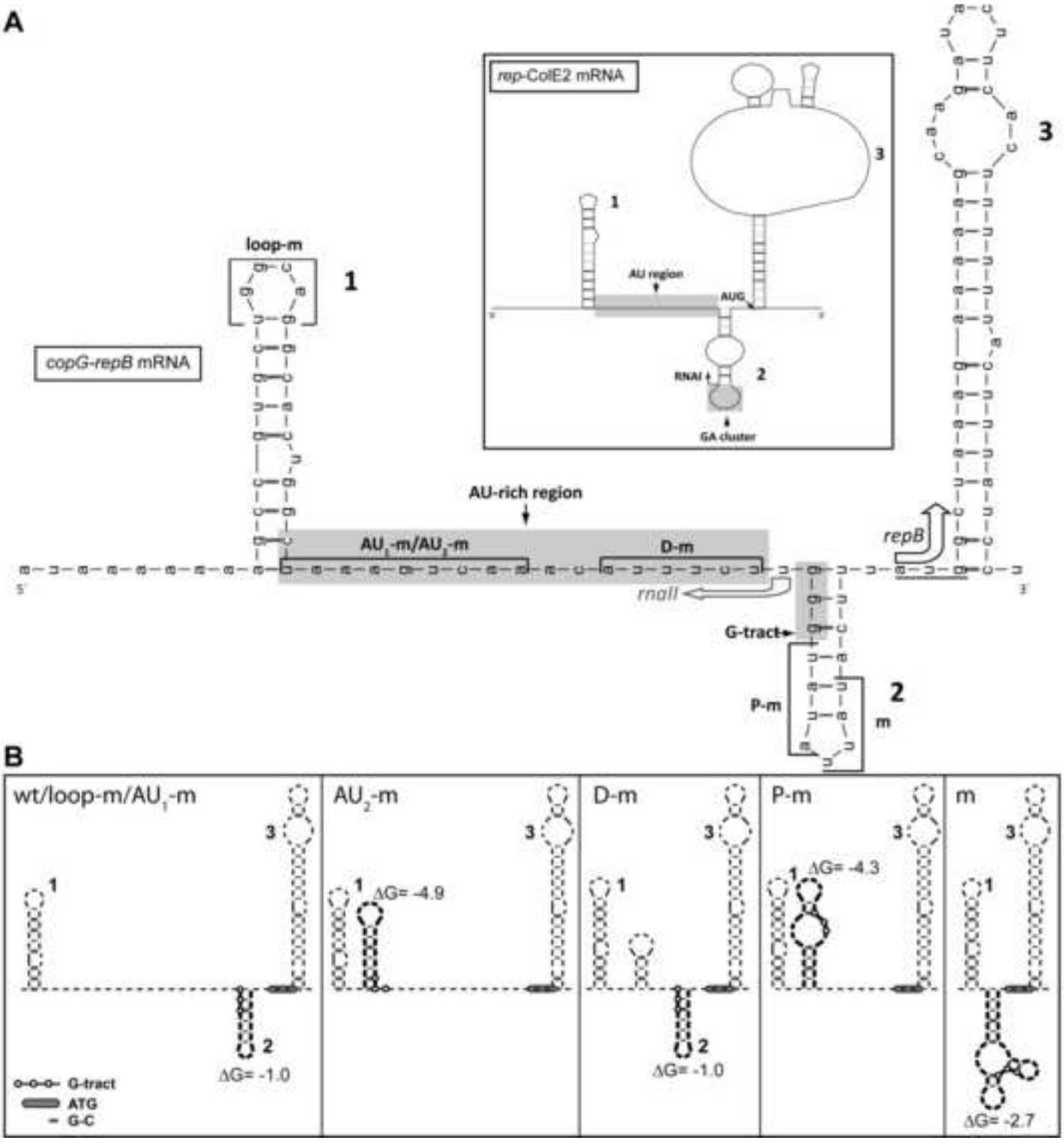

Figure 5

Click here to download high resolution image