Aminoacyl-tRNA synthetase genes of *Bacillus subtilis*: organization and regulation

Martin Pelchat and Jacques Lapointe

Abstract: In *Bacillus subtilis*, 14 of the 24 genes encoding aminoacyl-tRNA synthetases (aaRS) are regulated by tRNA-mediated antitermination in response to starvation for their cognate amino acid. Their transcripts have an untranslated leader mRNA of about 300 nucleotides, including alternative and mutually exclusive terminator–antiterminator structures, just upstream from the translation initiation site. Following antitermination, some of these transcripts are cleaved leaving at the 5′-end of the mature mRNAs, stable secondary structures that can protect them against degradation. Although most *B. subtilis* aaRS genes are expressed as monocistronic mRNAs, the *gltX* gene encoding the glutamyl-tRNA synthetase is cotranscribed with *cysE* and *cysS* encoding serine acetyltransferase and cysteinyl-tRNA synthetase, respectively. Transcription of *gltX* is not controlled by a tRNA, but tRNA_Cys-mediated antitermination regulates the elongation of transcription into *cysE* and *cysS*. The full-length *gltX-cysE-cysS* transcript is then cleaved into a monocistronic *gltX* mRNA and a *cysE-cysS* mRNA.

Key words: regulation, aminoacyl-tRNA synthetase, T-Box, processing.

Introduction

Protein biosynthesis plays a central role in the metabolism of all living organisms, consuming a major portion of their cellular energy (Ingraham et al. 1983). The numerous factors that catalyze this process include ribosomal proteins and RNAs, tRNAs, and aminoacyl-tRNA synthetases. In bacteria, the expression of these factors, encoded by more than 200 genes, must be efficient and coordinated. However, the mechanisms coordinating the expression of these genes and those involved in other metabolic pathways remain poorly understood. The aminoacyl-tRNA synthetases (aaRS) play an important role in protein biosynthesis by catalyzing the accurate charging of tRNAs with their cognate amino acids. As with most macromolecules involved in protein biosynthesis, the aaRS are generally present in high levels, and these levels increase with growth rate, a phenomenon known as metabolic regulation (Neidhardt et al. 1977).

Organisation of aaRS genes in *Bacillus subtilis*

The determination of the nucleotide sequence of the *Bacillus subtilis* genome (4 214 814 bp) by Kunst et al. (1997) allowed the identification of 24 aaRS genes specific for 19 amino acids (Putzer 1997). In this organism, there is no glutaminyl-tRNA synthetase, but a single glutamyl-tRNA synthetase is responsible for aminoacylating both tRNA\textsubscript{Glu} and tRNA\textsubscript{Glu} with glutamate (Lapointe et al. 1986; Pelchat et al. 1998); tRNA\textsubscript{Glu} is first charged with glutamate, which is subsequently converted to glutamine by an amidotransferase while it is attached to tRNA (Wilcox and Nirenberg 1968;
Curnow et al. 1997). Two heterodimeric aaRSs are encoded by distinct but cotranscribed genes: glycyI-tRNA synthetase by glyQS and glyS and phenylalanyl-tRNA synthetase bypheS and pheT. There are also two distinct aaRSs for each of the following amino acids: threonine (encoded by thrS and thrZ), tyrosine (encoded by tyrS and tyrZ) and histidine (encoded by hisS and hisZ; hisZ is known only from sequence comparison).

**Bacillus subtilis** aaRS genes are generally grouped on the chromosome. This is in contrast to what is found in *Escherichia coli*, where aaRS genes are scattered around the chromosome (Grunberg-Manago 1996; Blattner et al. 1997). In *B. subtilis*, half of these genes (glyQS, alaS, aspS, hisS, valS, pheS, tyrS, leuS) are between 223° and 268° (0°/360° being the replication origin), five (serS, metS, lysS, gltX, cysS) are near the replication origin between 2° and 10°, and three (argS, thrZ, tyrZ) are situated between 328° and 337°. Furthermore, all *B. subtilis* aaRS genes, except trpS and tyrZ, are in the same orientation as the replication fork, which is different from what is found in *E. coli* (Grunberg-Manago 1996; Blattner et al. 1997). This orientation is typical of highly expressed genes, probably because there is less interference between their transcription and genome replication (French 1992).

**Coexpression of aaRS genes**

In prokaryotes, the genes encoding enzymes involved in the same pathway are often cotranscribed. There have been few known cases of cotranscribed aaRS genes. Two of them are the gltX-cysE-cysS operons of *B. subtilis* and *Staphylococcus xylosus*, in which the genes encoding two aaRSs specific for glutamate and cysteine, respectively, are separated by a gene encoding the first enzyme of cysteine biosynthesis (Gagnon et al. 1994; Fiegler and Brückner 1997). Other reported cases are the *Rhizobium meliloti* glxL-lysS operon (Pelchat et al. 1999), in which two aaRS genes (specific for glutamate and lysine respectively) are separated by 10 bp, and the thrS-infC-rpmL-rplT-pheS-pheT-himA cluster of *E. coli* in which the genes encoding the aaRSs specific for threonine (thrS) and phenylalanine (pheS-pheT), respectively, are separated by three genes (Springer et al. 1985; Butler et al. 1986; Lesage et al. 1990; for a review, see Grunberg-Manago 1996).

**Transcriptional and posttranscriptional regulation of some *B. subtilis* aaRS genes**

In contrast to what is found in *E. coli*, where different mechanisms are used to regulate the expression of various aaRS genes (Grunberg-Manago 1996), in *B. subtilis* and other Gram-positive bacteria, many of these genes and many amino acid biosynthetic operons are regulated by a common mechanism. Among the 24 aaRS genes in *B. subtilis*, 14 are controlled by tRNA-mediated antitermination of transcription in response to starvation for their cognate amino acid (Henkin et al. 1992; Grundy and Henkin 1993; Putzer et al. 1992, 1995a). A highly conserved sequence, called the T-box, is situated just upstream from the terminator structure and can form, with the S-strand of the terminator stem, an alternative and mutually exclusive antitermination structure (Grundy and Henkin 1993). The much weaker antiterminator conformation is thought to be stabilized by base-pairing of the cognate uncharged tRNA with its NCCA-3′ end to a UGGN sequence in the T-box (Grundy et al. 1994; Putzer et al. 1995b) and with its anticodon to a specifier codon present in a loop near the 5′-end of the leader (Grundy and Henkin 1993; Luo et al. 1998) (Fig. 1). This model proposes that two contacts between uncharged tRNA and the leader may be sufficient to allow specific antitermination. When the regulatory tRNA, whose anticodon is complementary to the specifier codon is aminoacylated, it cannot interact with the T-Box, which allows the formation of the p-independent terminator structure and causes transcription termination. From this model, the over-expression of any aaRS having a T-Box leader in its mRNA should inhibit its gene expression via the aminoacylation of the regulatory tRNA. This is true for *B. subtilis* tyrS (Grundy and Henkin 1993), thrS, thrZ (Putzer et al. 1992), valS (Luo et al. 1997), and cysS (Gagnon et al. 1994), but not for pheST, where an overproduction of PheRS has no influence on the expression of a *pheS-lacZ* fusion (Putzer et al. 1995b). This observation and studies of *B. subtilis* thrS, thrZ, leuS, and valS suggest that the cognate aaRSs could be involved in the regulation of their genes by interacting directly with mRNA (for a review, see Putzer 1997).

Following this tRNA-mediated antitermination, the leaders of tyrS, serS, thrS, and ilv-leu mRNAs are cleaved between the T-Box consensus sequence and the terminator structure leaving a stable secondary structure at the 5′-end of the processed aaRS mRNA (Condon et al. 1996). For the *B. subtilis* thrS mRNA this processing increases its half-life and was reported to be catalyzed in vivo and in vitro by *E. coli* RNase E (Condon et al. 1996, 1997). Moreover, the ratio of the amounts of processed and full-length thrS transcripts is influenced by threonine starvation, indicating that this processing participates to thrS regulation.

**Processing of the *B. subtilis* gltX-cysE-cysS transcript**

In *B. subtilis*, the gltX gene is cotranscribed with cysE and cysS encoding, respectively, the serine acetyl-transferase (the first enzyme in the pathway for cysteine biosynthesis from serine) and the cysteyln-tRNA synthetase (Gagnon et al. 1994; Pelchat and Lapointe 1999). The mRNA of the intergenic region between gltX and cysE may adopt secondary structures similar to those found in the leader of other *B. subtilis* aaRS mRNAs (Fig. 1). We previously demonstrated that the expression of cysE and cysS is partly uncoupled from that of gltX by the presence of a RNA5′-dependent transcriptional attenuator between gltX and cysE (Gagnon et al. 1994). Moreover, the mRNA of the intergenic region between gltX, cysE and cysS is cleaved in vivo and in vitro immediately downstream from the p-independent terminator structure (Pelchat and Lapointe 1999), which differs by its position and its mechanism from that of the *B. subtilis* thrS transcript, reported to be catalyzed by *E. coli*.
RNase E in vivo and where no processing was observed in vitro in the absence of this enzyme (Condon et al. 1997).

The mechanism of this cleavage is not fully characterized, but the fact that it can be observed in vitro, as well as in vivo, suggests self-cleavage. However, we cannot rule out the possibility that this highly specific cleavage is due to the RNA polymerase (Surratt et al. 1991) or to the presence of one fragile phosphodiester bound in this long transcript (Watson et al. 1984; Hosaka et al. 1991; Kierzek 1992). This processing leaves stable putative secondary structures at the 3'-end of the gltX transcript and at the 5'-end of the cysE/S RNA (Fig. 2) (Pelchat and Lapointe 1999). The presence of these hairpins at the extremities of these processed mRNAs may protect them against degradation from their ends (see Emory et al. 1992), and is consistent with the observation of the longer half-life of the proximal part of cysE than that of gltX mRNA, which contains no such secondary structure in its short leader (Pelchat and Lapointe 1999).
This result also suggests that this specific cleavage of the gltX-cysE-cysS transcript may allow distinct regulation of these cotranscribed genes.

**Conclusion and future directions**

Among the 22 aaRS genes identified on the *B. subtilis* chromosome, 14 contain in their 5' flanking region a T-box and sequences whose identified or putative secondary structures suggest a common regulatory mechanism at the level of tRNA-dependent transcription antitermination (reviewed by Putzer 1997). Little is known about the regulation of the other eight aaRS genes, but it is likely that they are subject to metabolic regulation, as are all *E. coli* aaRS genes (Grunberg-Manago 1996). In the case of the *B. subtilis* gltX-cysE-cysS operon, it is unlikely that metabolic regulation could be mediated via a tRNA-dependent transcriptional regulation or via autoregulation because of the short length (about 45 nt) of the gltX 5'-leader (Gagnon et al. 1994). It is tempting to speculate that for gltX and maybe other members of this second group, metabolic regulation could depend on the energy charge of the cell via the intracellular levels of ATP and/or GTP, as shown for the promoters of *E. coli* tRNA genes (Gaal et al. 1997).

Cleavage of the gltX-cysE-cysS transcript takes place downstream of the p-independent terminator near the end of the transcript of the gltX-cysE intergenic region (Pelchat and Lapointe 1999), whereas it takes place upstream of this terminator in the other studied T-Box containing aaRS transcripts. This difference may allow a stabilization of the upstream gltX mRNA, and the presence of a putative hairpin immediately following this terminator appears to stabilize the downstream cysE-cysS transcript (Pelchat and Lapointe 1999). These observations suggest that these processing sites may have been selected by their stabilizing influence on the adjacent aminoacyl-tRNA synthetase mRNA and may have a general importance in the regulation of this family of genes.

**Acknowledgements**

This study was supported by grants from the Natural Sciences and Engineering Research Council of Canada (grant No. OGP0009597 to J.L.) and the Fonds pour la formation de chercheurs et l’aide à la recherche (FCAR; grant No. 97-ER-248) of Québec. M.P. was a postgraduate fellow of FCAR.

**References**


Lapointe, J., Duplain, L., and Proulx, M. 1986. A single glutamyl-tRNA synthetase aminoacylates tRNA^Glu_ and tRNA^Glu_ in *Bacillus*