

Co-transcription of *Rhizobium meliloti* lysyl-tRNA synthetase and glutamyl-tRNA synthetase genes

Martin Pelchat^a, Yves Gagnon^a, Serge Laberge^b, Jacques Lapointe^{a,*}

^aDépartement de Biochimie, Faculté des Sciences et de Génie, Pavillon Charles-Eugène Marchand, room 3210, Université Laval, Sainte-Foy, Que. G1K 7P4, Canada

^bStation de recherches sur les sols et les grandes cultures, Agriculture et Agroalimentaire Canada, Sainte-Foy, Que. G1V 2J6, Canada

Received 24 February 1999

Abstract An open reading frame encoding a putative polypeptide very similar to several lysyl-tRNA synthetases was found 10 nucleotides downstream of *Rhizobium meliloti* *gltX* encoding glutamyl-tRNA synthetase. Expression of this gene complemented a mutation in *lysS* of *Escherichia coli* and led to the overexpression of a polypeptide of the expected mass (62 kDa), thus confirming that it encodes *R. meliloti* lysyl-tRNA synthetase. Reverse transcription/polymerase chain reaction was used to demonstrate that this *lysS* gene is co-transcribed with *gltX* in *R. meliloti*. This is the first reported case of two immediately adjacent and co-transcribed genes encoding aminoacyl-tRNA synthetases.

© 1999 Federation of European Biochemical Societies.

Key words: Co-transcription; *Rhizobium meliloti* *gltX*; *Rhizobium meliloti* *lysS*; Glutamyl-tRNA synthetase; Lysyl-tRNA synthetase

1. Introduction

Aminoacyl-tRNA synthetases (aaRSs) constitute a family of enzymes, each of which catalyzes the formation of an ester linkage between a specific amino acid and the 3'-end ribose of its cognate tRNA(s) [1]. Each aaRS falls into one of two evolutionarily distinct classes according to their structures and catalytic properties [1–3]. The only exception to this rule is lysyl-tRNA synthetase (LysRS). LysRS of most organisms is a class II enzyme whereas that of several euryarchaea and *Borrelia burgdorferi* contains the amino acid sequence motifs characteristic of class I enzymes [4,5].

In bacterial genomes, there have been only three known cases of co-transcribed aaRS genes described to date. Two of them are the *gltX-cysE-cysS* operons of *Bacillus 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 [6,7]. The third case is the *thrS-infC-rpmL-rplT-pheS-pheT-himA* cluster of *Escherichia coli* in which the genes encoding the aaRSs specific for threonine (*thrS*) and phenylalanine (*pheS-pheT*), respectively, are separated by three genes [8–11]. Open reading frames (ORFs) apparently encoding the histidyl- and the aspartyl-tRNA synthetase are adjacent on the *B. subtilis* genome [12], but there has been no report about their co-transcription.

Two aaRS genes of *Rhizobium meliloti* have been cloned to date: *alaS*, which is located downstream of *recA* [13], and *gltX* [14]. We report here that a gene encoding a LysRS in

R. meliloti starts 10 nucleotides downstream of the UAA stop codon of *gltX*, and that it is co-transcribed with *gltX*. This is the first known operon containing immediately adjacent aaRS genes.

2. Materials and methods

2.1. Chemicals, enzymes and oligonucleotides

Restriction endonucleases and modification enzymes were purchased from Bethesda Research Laboratory Inc., Pharmacia LKB Biotechnology, Boehringer Mannheim, Sigma Chemical Co., New England Biolabs or International Biotechnologies, Inc. and were used according to the manufacturers' instructions. 5-Bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) and isopropyl-β-D-thiogalactopyranoside (IPTG) were obtained from IBI and BRL, respectively. The labeled nucleotides [α -³²P]dCTP (800 Ci/mmol) and [γ -³²P]ATP (3000 Ci/mmol) were obtained from DuPont Canada Inc. Double-stranded and single-stranded DNA were sequenced by the dideoxy-ribonucleotide chain termination technique [15], using the T7 sequencing kit from Pharmacia. The following oligonucleotides were synthesized: RHLYSS (5'-GGTCGAGCTTCTGGG-3'), RHLYS1 (5'-GACCGACAAGACACAGG-3') and RHGLTX (5'-CCATGGC-GGGCAGTG-3').

2.2. Bacterial cell growth and transformation

E. coli DH5 α (F[−] *recA1 endA1 gyrA96 thi-1 hsdR17* (R_K[−], M_K⁺) *supE44 relA1 λ*[−]) [16] was used to maintain and propagate plasmids. *E. coli* strains XA103 (F[−] Δ (*lac-pro*) *gyrA96 rpoB metB argE* (*Am*) *ara supF*) [17] and PAL3103KΔRS (XA103; Δ *lysS::kan*, *srl-300::Tn10 recA56*) [18] were used in complementation assays. *E. coli* and *R. meliloti* A2 were grown respectively on LB medium [19] and YMB medium [20,21]. Ampicillin was added to a final concentration of 100 μg/ml. *E. coli* cells were transformed according to Chung et al. [22]. Plasmid DNA was purified according to Brun et al. [23].

2.3. Plasmid construction

Chromosomal DNA was isolated from *R. meliloti* A2 as previously described [14]. Vector pTrc99b (Pharmacia) was used to express the putative *lysS* gene. This vector contains the IPTG-inducible *trc* promoter and the *lacI^q* gene, which increases the amount of *lac* repressor in the cell [24]. A homologous radiolabeled *EcoRV-SalI* DNA fragment was used as a Southern hybridization probe [19] to identify a 2.5 kb *Bam*HI fragment containing the putative *lysS* gene (Fig. 1), which was inserted into the unique *Bam*HI site of pTrc99b. The recombinant plasmids were mapped to determine the fragment orientation relative to the *trc* promoter. In pYGM305, the *trc* promoter and *lysS* are oriented in the same direction; this plasmid was digested to delete the upstream *Bam*HI-*StuI* region of the fragment which is not included in *lysS* (Fig. 1), to produce pYGM306. pYGM307 contains *lysS* inserted in the opposite orientation with respect to the *trc* promoter, and was used as control. Expression of proteins encoded by recombinant plasmids in *E. coli* was monitored as described by Amann et al. [24].

2.4. RT-PCR amplification of *lysS* mRNA

RNA was extracted from *R. meliloti* A2 according to Gray et al. [25]. mRNA sequences were amplified by PCR according to Ausubel et al. [19]. *R. meliloti* A2 total RNA (10 μg) was mixed with 4 pmol of

*Corresponding author. Fax: (1) (418) 656-3664.
E-mail: jacques.lapointe@bcm.ulaval.ca

Fig. 2. Intergenic sequence between *R. meliloti* *gltX* and *lysS*, and putative secondary structure of the 5' leader region of the *gltX-lysS* mRNA. A: The amino acid sequence deduced from the nucleotide sequence for the beginning of *gltX* and *lysS* is given in single-letter code. The putative promoter of *gltX* is shown in bold. The underlined nucleotides correspond to the stems, numbered from 1 to 4, of the putative cloverleaf-like secondary structure (B) found upstream of and near *gltX*.

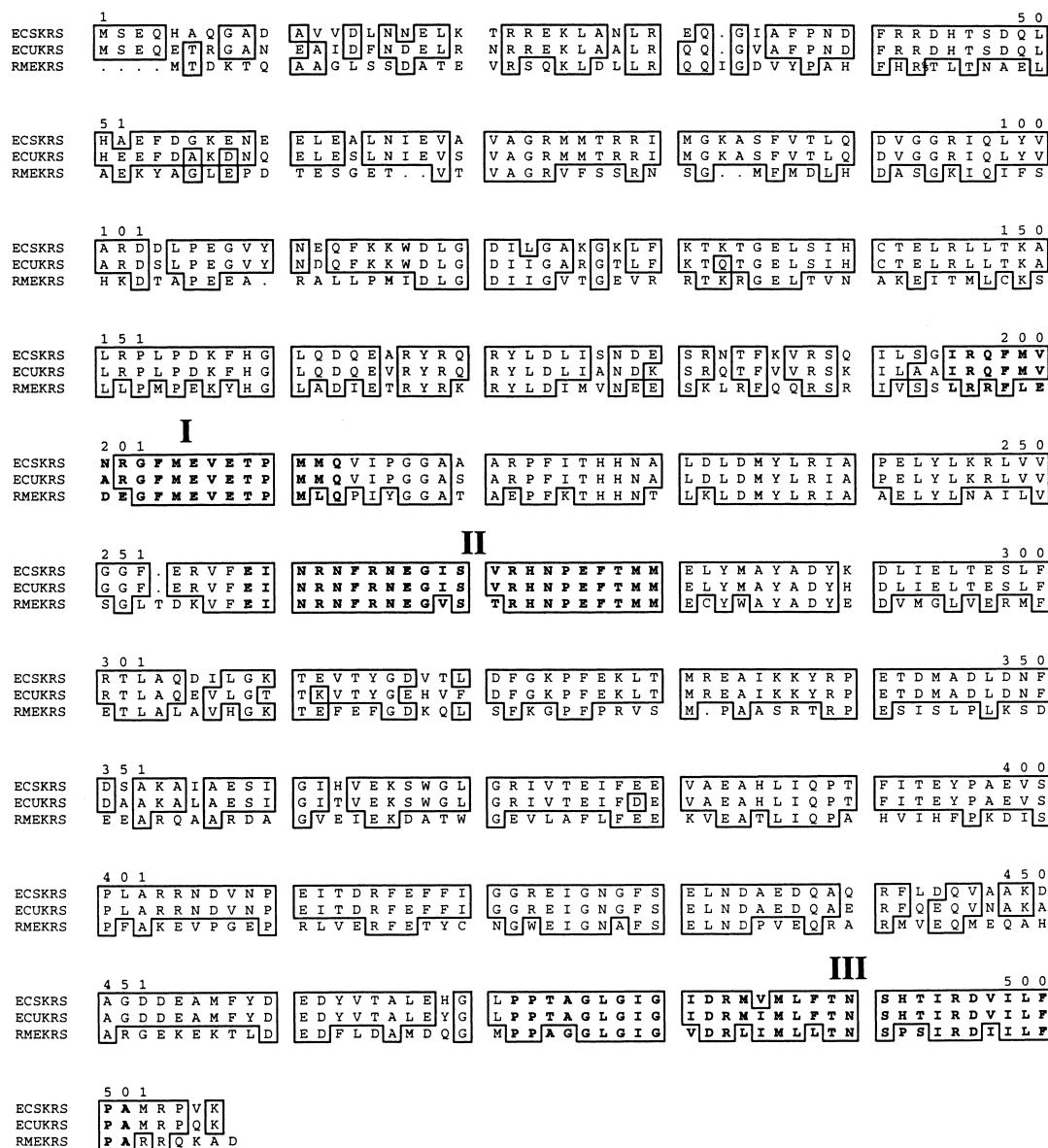


Fig. 3. Alignment of the deduced amino acid sequences of *R. meliloti* LysRS and those of the two *E. coli* LysRSs. The three conserved motifs (I, II and III) of the class II aaRSs are shown in bold. The sequences are identified as follows: *E. coli* *lysS*: ECSKRS [41]; *E. coli* *lysU*: ECUKRS [42]; *R. meliloti* *lysS*: RMEKRS [42].

leaf-like secondary structure ($\Delta G = -19.7$ kcal/mol) [27] (Fig. 2B).

Computer analysis of the deduced amino acid sequence of the protein encoded by the *R. meliloti* ORF showed higher similarities with the two *E. coli* LysRSs than with any other aaRS (Fig. 3). Moreover, the predicted sequence of this protein is about 47% identical to that of the two *E. coli* LysRSs and less than 30% identical to those of AspRS and AsnRS from *E. coli*, which together with LysRS constitute a subclass of class II aaRSs [3,28]. The three conserved motifs unique to class II aaRSs are present in the *R. meliloti* ORF. Although these motifs are known to be involved in the catalytic activity [2], their respective sequences are not highly conserved between aaRSs of different specificities [3]. However, the respective sequences of these motifs are highly conserved in aaRSs of the same specificity, enabling their use in identifying unknown class II aaRSs. The fact that the sequences of the three

motifs found in the protein encoded by the *R. meliloti* ORF immediately downstream of *gltX* show the highest identity ($\sim 70\%$) with those of LysRSs suggests that this ORF encodes a LysRS.

3.2. Expression of the putative *R. meliloti* *lysS* gene in *E. coli* and complementation of a *lysS* mutation

To demonstrate that the ORF downstream of *R. meliloti* *gltX* encodes a functional LysRS, it was expressed in *E. coli* and its ability to complement a mutant of the *lysS* gene was verified. Since no promoter was found in front of the putative *R. meliloti* *lysS*, this gene was expressed in *E. coli* using the inducible *trc* promoter. Two plasmids, pYGM306 and pYGM307, were constructed (see Section 2) and introduced into *E. coli*. The protein content of strain PAL3103KARS bearing different plasmids was examined by Coomassie blue-stained SDS polyacrylamide gels (Fig. 4). A band correspond-

ing to a polypeptide of apparent molecular mass 62 kDa was detected in IPTG-induced cells containing pYGM306, which contains the insert in the right orientation (Fig. 4, lane 1). This value is consistent with the molecular mass calculated from the deduced amino acid sequence of LysRS (62 907 Da). The same band was not detected in cells containing pYGM307, in which *lysS* is inserted in the wrong orientation (Fig. 4, lanes 3 and 4). Finally, the band was also detected in cells containing pYGM306 even when they were not induced with IPTG (Fig. 4, lane 2), suggesting that the control of the *trc* promoter was leaky.

The function of the putative *R. meliloti lysS* gene was tested by complementation assays using *E. coli* strain PAL3103KΔRS. This strain has a cold-sensitive phenotype resulting from the disruption of *lysS* and grows very slowly at 25°C [18]. The slow growth of this strain was complemented efficiently by the recombinant plasmid pYGM306 in the presence of 1 mM IPTG, and only partially in its absence. In contrast, this phenotype was not complemented by pYGM307. These results show that the ORF immediately downstream of *gltX* in *R. meliloti* encodes a functional LysRS.

3.3. Co-transcription of *gltX* and *lysS* in *R. meliloti*

The initiation codon (AUG) of *lysS* is situated 10 nt downstream of the termination codon (UAA) of *gltX*, suggesting that these two genes are co-transcribed. To verify this co-regulation, the RNA extracted from *R. meliloti* A2 was reverse-transcribed using RHLYSS, an oligonucleotide situated 70 nt downstream of the beginning of *lysS* (Fig. 1). Two segments of the corresponding cDNA were amplified by the polymerase chain reaction, using RHLYSS and either an oligonucleotide situated at the beginning of *lysS* (RHLYS1) or an oligonucleotide situated near the end of *gltX* (RHGLTX) (Fig. 5A). The positions of the sequences corresponding to these oligonucleotides are indicated in Fig. 1. The identity of these amplicons was confirmed by their hybridization with radiolabeled RHLYS1 (Fig. 5B). The resulting amplicons of the *gltX-lysS* and *lysS* regions were of the expected size of 113 nt and 68 nt, respectively (Fig. 5, lanes 1 and 3). These amplicons were not observed when the reverse transcription reaction was omitted (Fig. 5, lanes 2 and 4), indicating that they did not result from the amplification of contaminating chromosomal DNA in the purified RNA preparation. These results show that *gltX* and *lysS* are co-transcribed in vivo. For unknown reasons, in vivo transcripts of *gltX* and/or *lysS*

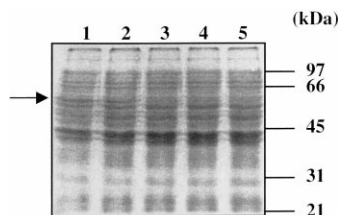


Fig. 4. Expression of the *R. meliloti lysS* gene in *E. coli*. Total protein extracts of *E. coli* PAL3103KΔRS harboring the indicated plasmids were separated by SDS-PAGE on a 10% gel. Lanes 1 and 2: PAL3103KΔRS bearing plasmid pYGM306. Lanes 3 and 4: PAL3103KΔRS bearing plasmid pYGM307. The extracts analyzed in lanes 1 and 3 were from cells induced with 1 mM IPTG. Lane 5: PAL3103KΔRS without plasmid. The arrow indicates the position of a 62 kDa polypeptide.

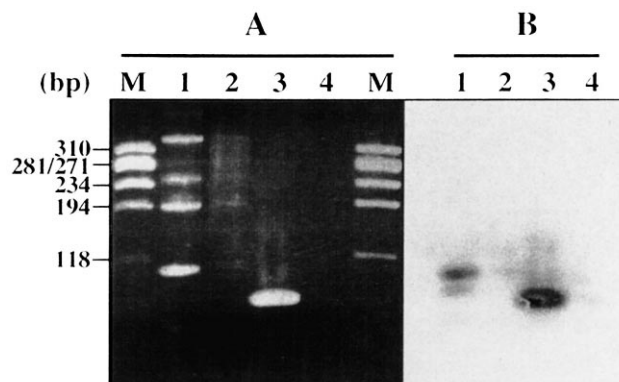


Fig. 5. Detection by RT-PCR of a transcript overlapping *R. meliloti gltX* and *lysS*. A: cDNA obtained by reverse transcription of total *R. meliloti* A2 RNA was amplified using an oligonucleotide situated 70 nt after the beginning of *lysS* (RHLYSS) and another one situated either at the end of *gltX* (RHGLTX; lane 1) or at the beginning of *lysS* (RHLYS1; lane 3) (see Fig. 1). Controls were performed using the same oligonucleotides on RNA prior to reverse transcription (lanes 2 and 4). B: Samples were transferred onto a nylon membrane and hybridized with radiolabeled RHLYS1 to confirm their identity. M, *Hae*III-digested ϕ X174 DNA.

could not be detected by Northern hybridization or primer extension using several different primers.

4. Discussion

We report the identification, cloning, and the complete nucleotide sequence of the *R. meliloti lysS* gene encoding the lysyl-tRNA synthetase and its co-transcription with the *gltX* gene encoding the glutamyl-tRNA synthetase. The deduced amino acid sequence of the *R. meliloti* LysRS contains the three motifs conserved among the class II aaRSs and shows strong similarity to the two LysRSs of *E. coli*. The overexpression of the *R. meliloti lysS* gene in *E. coli* functionally complemented a mutation in the latter's *lysS* and was correlated with the presence of a polypeptide of 62 kDa.

The 5'-flanking region of *R. meliloti gltX* contains two hexamers, TTGACA-15 nt-GG-TCTCAC-G (Fig. 2A), which have a total of nine identities with the consensus sequence of the -35/-10 promoters of *E. coli* (TTGACA-17 nt-TA-TAAT) [29]. The two hexamers also have several identities with the two characterized vegetative promoters of *R. meliloti* (TTGACC-15 nt-GG-GAGAAAT-G) [30]. This putative promoter of the *gltX-lysS* operon would specify the transcription of a 5' leader of about 125 nt, most of which can be folded in a putative cloverleaf-like secondary structure ($\Delta G = -19.7$ kcal/mol) (Fig. 2B). It is known that mRNA secondary structure influences the expression of *E. coli thrS* and of many Gram-positive aaRSs genes [8,31]. In *E. coli*, the 5' leader of *gltX* mRNA can adopt a stable secondary structure (close to the ribosome binding site) that is similar to the anticodon and part of the dihydroU stem and loop of tRNA^{Glu}, and could be implicated in mRNA processing upstream of the *E. coli gltX* gene [32]. The function of the putative secondary structure upstream of the *R. meliloti gltX-lysS* bicistronic mRNA has yet to be determined; it could be involved in its stabilization, as shown for secondary structures at the 5' end of other mRNAs [33], or it could be a recognition signal for specific RNases.

Several bacterial aaRS genes are adjacent to and, at least in some cases, probably co-regulated with other protein biosynthesis genes [11]; for instance, in several organisms, the promoter upstream of *gluX* also drives the transcription of another protein biosynthesis gene. In *B. subtilis* and *S. xylosus*, the *gluX* gene is co-transcribed with *cysE* and *cysS* encoding the serine acetyltransferase and the cysteinyl-tRNA synthetase respectively [6,7]. However, the expression of *cysE* and *cysS* is partly uncoupled from that of *gluX* by the presence of a tRNA^{Cys}-dependent transcriptional attenuator between *gluX* and *cysE* and by a specific cleavage of the long *gluX-cysE-cysS* transcript in the same region [34]. In higher eukaryotes, the largest polypeptide of the multisynthetase complex is a multifunctional aminoacyl-tRNA synthetase endowed with glutamyl-tRNA synthetase and prolyl-tRNA synthetase activities in its N-terminal and C-terminal halves, respectively [35]. A low molecular weight form of prolyl-tRNA synthetase has also been found in *Drosophila* and in rat salivary glands. Messenger RNAs encoding the full length glutamyl-prolyl-tRNA synthetase and just the prolyl-tRNA synthetase are controlled by two independent promoters in *Drosophila* [36,37]. It was recently observed that the initiation codon of *Azospirillum brasilense cysS* gene is situated 34 nt from the termination codon of *gluX* [38], but the possibility that they are co-transcribed was not tested. Furthermore, *gluX* in many organisms is adjacent to a tRNA operon [32,39,40]. These gene arrangements may have been selected to optimize protein biosynthesis.

Acknowledgements: This work was supported by Grant OGP0009597 from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Grant 2227 from CORPAQ to J.L. M.P. was a postgraduate fellow of the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR). Y.G. was a postgraduate fellow of NSERC and FCAR. We thank Dr. Lindsay D. Eltis for his judicious comments on the manuscript, and Dr. Lucien Bordeleau for providing *R. meliloti* A2. This is contribution #606 from the Station de recherches sur les sols et les grandes cultures, Agriculture et Agroalimentaire Canada, Ste-Foy, Que., Canada.

References

- [1] Meinel, T., Mechulam, Y. and Blanquet, S. (1995) in: tRNA: Structure, Biosynthesis, and Function (Söll, D. and RajBhandary, U.L., Eds.), pp. 251–333, ASM Press, Washington, DC.
- [2] Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) *Nature* 347, 203–206.
- [3] Moras, D. (1992) *Trends Biochem. Sci.* 17, 159–164.
- [4] Ibba, M., Curnow, A.W. and Söll, D. (1997) *Trends Biochem. Sci.* 22, 39–42.
- [5] Ibba, M., Morgan, S., Curnow, A.W., Pridmore, D.R., Vothknecht, U.C., Gardner, W., Lin, W., Woese, C.R. and Söll, D. (1997) *Science* 278, 1119–1122.
- [6] Gagnon, Y., Breton, R., Putzer, H., Pelchat, M., Grunberg-Manago, M. and Lapointe, J. (1994) *J. Biol. Chem.* 269, 7473–7482.
- [7] Fiegler, H. and Brückner, R. (1997) *FEMS Microbiol. Lett.* 148, 181–187.
- [8] Springer, M., Mayaux, J.F., Fayat, G., Plumbridge, J.A., Graffe, M., Blanquet, S. and Grunberg-Manago, M. (1985) *J. Mol. Biol.* 181, 467–478.
- [9] Butler, J.S., Springer, M., Dondon, J. and Grunberg-Manago, M. (1986) *J. Bacteriol.* 165, 198–203.
- [10] Lesage, P., Truong, H.N., Graffe, M., Dondon, J. and Springer, M. (1990) *J. Mol. Biol.* 213, 465–475.
- [11] Grunberg-Manago, M. (1996) in: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology* (Neidhardt, F.C., Ed.), pp. 1432–1457, ASM Press, Washington, DC.
- [12] Kunst, F. et al. (1997) *Nature* 390, 249–256.
- [13] Selbitschka, W., Arnold, W., Priefer, U.B., Rottschaefer, T., Schmidt, M., Simon, R. and Puhler, A. (1991) *Mol. Gen. Genet.* 229, 86–95.
- [14] Laberge, S., Gagnon, Y., Bordeleau, L.M. and Lapointe, J. (1989) *J. Bacteriol.* 171, 3926–3932.
- [15] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- [16] Hanahan, D. (1983) *J. Mol. Biol.* 166, 557–580.
- [17] Coulondre, C. and Miller, J.H. (1977) *J. Mol. Biol.* 117, 525–567.
- [18] Lévêque, F., Gazeau, M., Fromant, M., Blanquet, S. and Plateau, P. (1991) *J. Bacteriol.* 173, 7903–7910.
- [19] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) in: *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley-Interscience, New York.
- [20] Vincent, J.M. (1970) in: *A Manual for the Practical Study of the Root Nodule Bacteria*, IBP (International Biological Programme) Handbook No. 15, Blackwell Scientific, Oxford.
- [21] Bordeleau, L.M., Antoun, H. and Lachance, R.A. (1977) *Can. J. Plant Sci.* 57, 433–439.
- [22] Chung, C.T., Niemela, S.L. and Miller, R.H. (1989) *Proc. Natl. Acad. Sci. USA* 86, 8275–8279.
- [23] Brun, Y.V., Breton, R. and Lapointe, J. (1991) *DNA Seq.* 1, 285–289.
- [24] Amann, E., Ochs, B. and Abel, K.J. (1988) *Gene* 69, 301–315.
- [25] Gray, J.X., Djordjevic, M.A. and Rolfe, B.G. (1990) *J. Bacteriol.* 172, 193–203.
- [26] Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387–395.
- [27] Walter, A.E., Turner, D.H., Kim, J., Lyttle, M.H., Muller, P., Mathews, D.H. and Zuker, M. (1994) *Proc. Natl. Acad. Sci. USA* 91, 9218–9222.
- [28] Cusack, S., Härtlein, M. and Leberman, R. (1990) *Nucleic Acids Res.* 19, 3489–3498.
- [29] Neidhardt, F.C. and Van Bogelen, R.A. (1996) in: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology* (Neidhardt, F.C., Ed.), 2nd edn., ASM Press, Washington, DC.
- [30] Leong, S.A., Williams, P.H. and Ditta, G.S. (1985) *Nucleic Acids Res.* 13, 5965–5976.
- [31] Henkin, T.M. (1994) *Mol. Microbiol.* 13, 381–387.
- [32] Brun, Y.V., Sanfacon, H., Breton, R. and Lapointe, J. (1990) *J. Mol. Biol.* 214, 845–864.
- [33] Belasco, J.G. and Higgins, C.F. (1988) *Gene* 72, 15–23.
- [34] Pelchat, M. and Lapointe, J. (1999) *RNA* 5, 1–9.
- [35] Cerini, C., Kerjan, P., Astier, M., Gratecos, D., Mirande, M. and Sémériva, M. (1991) *EMBO J.* 10, 4267–4277.
- [36] Cerini, C., Sémériva, M. and Gratecos, D. (1997) *Eur. J. Biochem.* 244, 176–185.
- [37] Ting, S.M., Bogner, P. and Dignam, J.D. (1992) *J. Biol. Chem.* 267, 17701–17709.
- [38] Zimmer, W., Wesche, M. and Timmermans, L. (1998) *Curr. Microbiol.* 36, 327–331.
- [39] Eiglmeier, K., Honore, N., Woods, S.A., Caudron, B. and Cole, S.T. (1993) *Mol. Microbiol.* 7, 197–206.
- [40] Fleischmann, R.D. et al. (1995) *Science* 269, 496–512.
- [41] Kawakami, K., Jönsson, Y.H., Björk, G.R., Ikeda, H. and Nakamura, Y. (1988) *Proc. Natl. Acad. Sci. USA* 85, 5620–5624.
- [42] Lévêque, F., Plateau, P., Dessen, P. and Blanquet, S. (1990) *Nucleic Acids Res.* 18, 305–312.