JOURNAL OF BACTERIOLOGY, Dec. 1994, p. 7387–7390
0021-9193/94/$04.00+0

Vol. 176, No. 23

Characterization of the *Thermus thermophilus* Locus Encoding Peptide Deformylase and Methionyl-tRNA$_f^{Met}$ Formyltransferase

THIERRY MEINNEL* AND SYLVAIN BLANQUET

Laboratoire de Biochimie, Unité de Recherche Associée no. 240 du Centre National de la Recherche Scientifique, Ecole Polytechnique, F-91128 Palaiseau Cedex, France

Received 28 June 1994/Accepted 20 September 1994

An *Escherichia coli* strain with thermosensitive expression of the gene encoding peptide deformylase (*fms*) has been constructed. At nonpermissive temperatures, this strain fails to grow. The essential character of the *fms* gene was further used to clone by heterologous complementation the locus corresponding to *Thermus thermophilus* peptide deformylase. The cloned fragment also carries the methionyl-tRNA$_f^{Met}$ formyltransferase gene (*fmt*). It is located immediately downstream from the *fms* gene, as in *E. coli*. Further sequence analysis of the region surrounding the *E. coli fms-fmt* locus indicates that the genes bordering the *fms-fmt* region are not conserved in *T. thermophilus*.

Methionine, the universal start signal in mRNA translation, undergoes a series of transformations prior to its incorporation into proteins (reviewed in reference 16). Thereafter, it is removed in most cases from mature proteins by a ubiquitous enzyme, methionine aminopeptidase. In procaryotes and most likely in the organelles of eucaryotes, because of the N-formylation of methionyl-tRNA$_f^{Met}$, the removal of methionine from nascent proteins first requires the action of a peptide deformylase (PDF), and methionine aminopeptidase acts once the formyl group has been removed (1, 12, 23). All the *Escherichia coli* genes belonging to the so-called methionine pathway in translation initiation are now cloned (see references in reference 16). Only the genes encoding peptide deformylase (*fms* [6, 14], also called *def* [21]) and methionyl-tRNA$_f^{Met}$ formyltransferase (MTF) (*fmt* [6]) are physically related and belong to the same transcriptional unit (15). Interestingly, since the products of these two genes are very closely related at the functional level, such a genetic linkage in *E. coli* appears to be particularly appropriate. In this context, it may be expected that, in other procaryotes, the genes corresponding to the two MTF and PDF activities could also be associated in the same transcriptional unit. Also, since our recent biochemical study indicated that PDF from *E. coli* might be related to the family of zinc proteases (14), it was useful to search for the occurrence of the amino acid motif characteristic of the above family, HEXXH (10, 24), in the PDF from another procaryote.

The present work deals with the cloning and the sequencing of the *Thermus thermophilus* gene encoding PDF. The cloning strategy was based on the idea that PDF activity should be important for *E. coli* growth, at least because, beyond the deformylation step, methionine aminopeptidase is strictly required for *E. coli* and *Salmonella typhimurium* cell growth (3, 18). Since methionine aminopeptidase cannot cleave N-blocked methionine peptides (see references in reference 16), an absolute requirement for a PDF activity was therefore expected. An *E. coli* strain whose growth strictly depended on a thermosensitive expression of PDF was constructed and used to clone the *T. thermophilus fms-fmt* locus by genetic complementation.

* Corresponding author. Phone: 33 1 69 33 48 80. Fax: 33 1 69 33 30 13. Electronic mail address: labo@coli.polytechnique.fr.

Construction of a conditional lethal strain of the *fms* gene. The *E. coli* strains and plasmids used in this study are shown in Table 1. In *E. coli*, *fmt* and *fms* are transcribed from the same promoter, with *fms* upstream from *fmt* (15). Since *fmt* is necessary for optimal cell growth (6), the expression of the *fmt* gene must be preserved when inactivating the only *fms* gene. Actually, the doubling time of the bacterium becomes significantly increased when the intracellular concentration of MTF is lowered by up to 10-fold (6). To evaluate the consequence on the *fmt* gene expression of an inactivation of the *fms* gene, a 343-bp *Eco*RV-*Sma*I deletion within *fms* only was created in plasmid pBS936XB. This yielded plasmid pBS936XBΔES. PDF and MTF activities in crude extracts of JM101Tr cells (8) carrying plasmid pBS936XB or pBS936XBΔES were compared. As expected, the PDF activity was no longer overexpressed from plasmid pBS936XBΔES. In turn, MTF activity was decreased fivefold in the extract of JM101Tr-pBS936XBΔES cells compared with that in the JM101Tr-pBS936XB extract.

Assuming that the replacement on the chromosome of the *fms-fmt* locus by its copy with the *Eco*RV-*Sma*I deletion should lower the intracellular formylase activity by a similar factor, the specific inactivation of the chromosomal *fms* gene was carried out. The *Hin*dII-*Hin*dII fragment of plasmid pEform (Fig. 1) (15) was first deleted to make the *Eco*RV site of *fms* unique (Fig. 1A). In a second step, the *Eco*RV-*Sma*I deletion in the above constructed plasmid yielded plasmid pEformΔ*fms*. We then inserted the *Kpn*I-*Bam*HI fragment of pEformΔ*fms* between the same sites of pMAK705, a plasmid which carries a thermosensitive replicon (7). With the help of the resulting plasmid (pMAKΔ*fms* [Fig. 1A]) and using the procedure described by Hamilton et al. (7), we constructed from *E. coli* K37 (19) strain PAL421-pMAKfms with the inactivated *fms* gene on the chromosome and the wild-type allele on the pMAK vector. PAL421-pMAKfms was then made recombination deficient by conjugation with the *recA56* Hfr strain JC10240 (yielding Pal421Tr-pMAKfms). At the nonpermissive temperature (42°C), the above constructed strain failed to grow on Luria-Bertani (LB) plates, indicating that an active *fms* gene was needed for cell doubling. At 37°C, the strain grew. It was however verified that plasmid pMAKfms had been retained in the cells. To show that the observed lethality at 42°C was caused by the absence of an active PDF gene, strain Pal421Tr-pMAKΔfms was transformed with pUCdef, a pUC18 derivative expressing PDF under the control of the *lac* pro-

TABLE 1. *E. coli* strains and plasmids used in this study

Strain or plasmid	Genotype or markers	Reference
Strains		
PAL13Tr	Derivative of JM101Tr; *fmtΔ1::kan*	6
JC10240	Hfr(PO45) *thr-3000 recA56 srl-300*::Tn*10 relA ilv-318 spoT1 thi-1 rpsE2300*	4
PAL421	*galK rpsL fmsΔ1*	This work
PAL421Tr	Derivative of PAL421; *recA56 srl-300*::Tn*10*	This work
Plasmids		
pMAK705	*cat*; thermosensitive pSC101 replicon	7
pBS936	*bla fms fmt* (contains an 8-kb chromosomal insertion of the *fms-fmt* region)	6
pEform	*bla fms fmt'*, derivative of pEBNB (15) with the *Bam*HI-*Nco*I fragment of pBS936	15 (Fig. 2)
pEformΔfms	*bla fmt*; derivative of pEform by deletion of the *Eco*RV-*Sma*I fragment	This work
pMAKΔfms	Derivative of pMAK705 by insertion of the *Kpn*I-*Bam*HI fragment of pEformΔfms	(Fig. 1)
pMAKfms	Derivative of pMAKΔfms by double crossing-over with the chromosomal DNA of K37; *fms*	This work
pUCTT5	Derivative of pUC18 by insertion of a 2.3-kb fragment within the *Bam*HI site; *fms fmt*	This work (Fig. 2)
pUCTT6	Derivative of pUC18 by insertion of a 6.5-kb fragment within the *Bam*HI site; *fms fmt*	This work (Fig. 2)
pMAF	Derivative of pMAK705; *fmt*	6

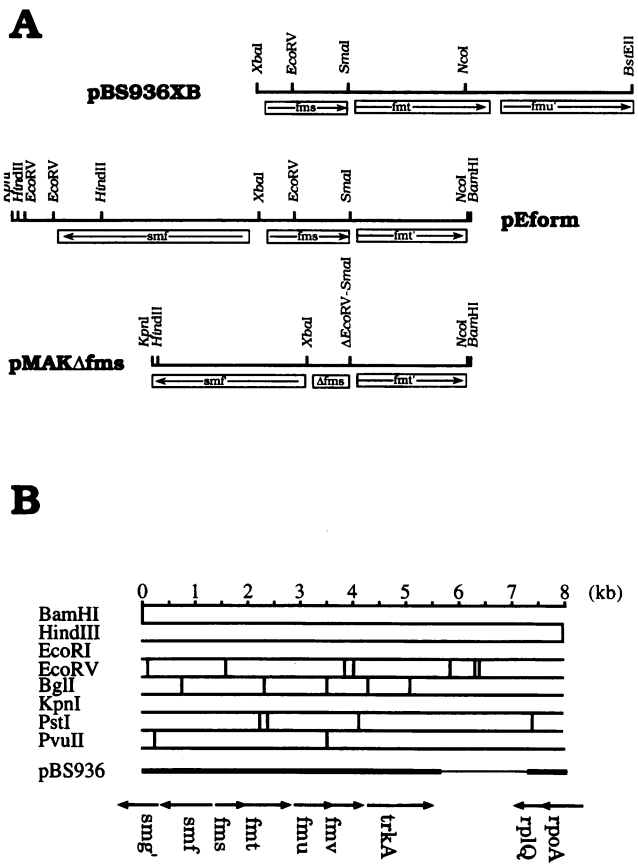

FIG. 1. Restriction maps of plasmids pBS936XB, pEform, and pMAKΔfms and of the region surrounding the *E. coli fms-fmt* locus on the pBS936 plasmid. (A) Restriction maps of plasmids pBS936XB, pEform, and pMAKΔfms. The *Xba*I-*Bst*EII region of plasmid pBS936XB (6) and the *Kpn*I-*Bam*HI region of plasmids pEform (15) and pMAKΔfms (this study) are shown. The main restriction sites as well as the localization of the ORFs covering the regions are indicated. The length of the *Kpn*I-*Bam*HI fragment of pEform is approximately 2,700 bp. (B) Physical map of the region surrounding the *E. coli fms-fmt* locus on the pBS936 plasmid. The occurrences of the indicated restriction sites are indicated by vertical bars. The scale is labelled on top. With the exception of the *Eco*RV sites, the restriction map of this region matches that of bp 3452 to 3460 of the *E. coli* chromosome, corresponding to λ phage 629 (11). Note that the location of the *fms-fmt* locus on the chromosomal insertion of λ phage 629 was checked by Southern blot analysis (not shown). The two parts of known nucleotide sequence of plasmid pBS936 are indicated in boldface and are available with EMBL accession numbers X77091 (this work) and X02543 (2).

moter (14). Contrarily to PAL421Tr-pUC18, strain PAL421Tr-pUCdef grew normally at 42°C on LB plates. From these results, we concluded that the *fms* gene was essential for *E. coli* viability.

Cloning and nucleotide sequencing of the *fms-fmt* locus from *T. thermophilus*. A DNA library from strain VK-1 was constructed by limited hydrolysis of chromosomal DNA with restriction endonuclease *Sau*3A. Fragments of about 2 to 10 kb were inserted in plasmid pUC18 previously hydrolyzed with *Bam*HI. Ligation products were used to transform the above constructed thermosensitive strain, PAL421Tr-pMAKfms. About 50,000 transformants were plated at 42°C on LB medium supplemented with ampicillin and isopropyl-β-D-thiogalactopyranoside (IPTG). After a 24-h incubation, two colonies could be observed. DNAs of the two corresponding plasmids (pUCTT5 and pUCTT6) were prepared and used to retransform strain PAL421Tr-pMAKfms. Growth at 42°C of the two resulting strains established that the thermoresistance phenotype was indeed linked to the occurrence of either plasmid. In both cases, we observed that growth depended on the addition of IPTG in the culture medium. To find out whether the locus encoding MTF could also be close to that encoding PDF in *T. thermophilus*, strain PAL13Tr-pMAF, which displays a MTF(Ts) phenotype, was transformed with either pUCTT5 or pUCTT6. In both cases, the plasmid-bearing strains became thermoresistant in the presence of IPTG, thereby indicating that the plasmids contained the gene for MTF, in addition to that for PDF. Restriction analysis of the 2.3- and 6.5-kb chromosomal insertions of the two plasmid DNAs pUCTT5 and pUCTT6 indicated that they were overlapping and that they shared the same 5′ cloning site downstream from the *lacZ* promoter (Fig. 2).

The nucleotide sequence of the 2.3-kb chromosomal insertion of plasmid pUCTT5 (available with GenBank and EMBL accession number X79087) revealed four open reading frames (ORFs) (Fig. 2). Codon usage of these 4 ORFs perfectly matched the one calculated from the 58 ORFs from *T. thermophilus* available in the NBRF protein data bank (release number 38). The four ORFs had the same orientation and were under the control of the *lacZ* promoter of pUC18 (Fig. 2). Only the two central ORFs (Orf2 and Orf3) were entirely borne on plasmid pUCTT5. The C terminus of Orf4 and the N

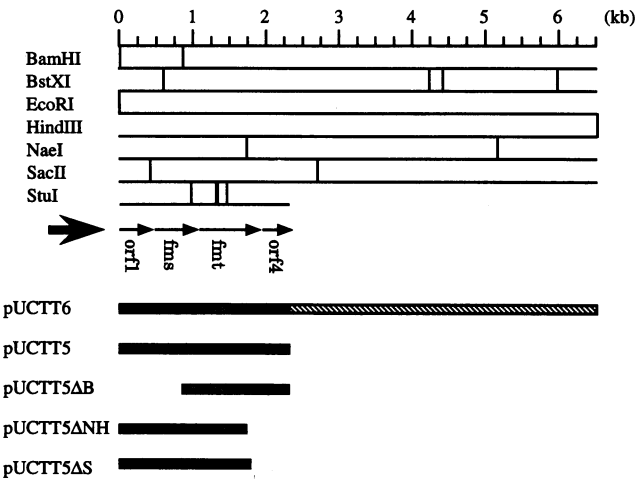

FIG. 2. Physical map of plasmids pUCTT5 and pUCTT6 and of derivatives of pUCTT5. The *Eco*RI-*Hin*dIII regions of plasmids pUCTT5 and pUCTT6 are shown. The scale is labelled on top. Restriction sites *Bam*HI, *Bst*XI, *Eco*RI, *Hin*dIII, *Nae*I, and *Sac*II are shown on both maps by vertical bars. The *Stu*I sites are also indicated in the case of pUCTT5. The locations of the four ORFs are indicated by thin arrows, and the direction from the *lacZ* promoter is shown by a thick one. Plasmid derivatives of pUCTT5 are shown on the bottom.

terminus of Orf1 were missing. Deletions within plasmid pUCTT5 were created to map which ORF encoded the *fms* and *fmt* genes of *T. thermophilus* (Fig. 2). The *fmt*(Ts) strain could be complemented at 42°C by a plasmid with a 5′ *Bam*HI-*Bam*HI deletion, whereas the *fms*(Ts) strain could not. Contrarily, a plasmid with a 3′ deletion (*Nae*I-*Hin*dII) allowed only the growth of the *fms*(Ts) strain at the nonpermissive temperature. Finally, an internal *Stu*I-*Stu*I deletion within pUCTT5 inactivated both genes. These experiments unambiguously identified Orf2 and Orf3 as the *fms* and *fmt* genes, respectively. Interestingly, the ATG translation start site of *fmt* overlaps the TGA translation stop site of *fms*. This organization strongly supports the idea that the two *T. thermophilus* genes might be expressed from the same promoter, as already established for *E. coli* (15). Moreover, the fact that the expression of the two activities depends on the addition of IPTG shows that in the *E. coli* context the two loci are cotranscribed from the *lacZ* promoter, i.e., strong transcriptional termination does not occur between the two genes. All these observations strongly support the idea that *fms* and *fmt* are part of the same transcriptional unit in the thermophilic bacterium.

With the help of the FASTA program (22), we searched in protein data banks (NBRF release 38, Swissprot release 28, and Yeast library MIPS) for proteins sharing sequence identities with the amino acid sequences of the four above-identified ORFs. As could be expected, the best match of the *T. thermophilus fms* and *fmt* gene products was observed with the corresponding *E. coli fms* and *fmt* gene products, with 35.6 and 41.6% amino acid identity, respectively (Fig. 3A and B). In particular, the HEXXH motif of zinc metalloproteases originally found in the *E. coli* PDF (14) is strictly conserved in the *T. thermophilus* enzyme. Moreover, the H(G or P)SLLPX(H, F, W, or Y)XG motif characteristic of formyltetrahydrofolate-binding proteins (16) occurred in the *T. thermophilus* MTF sequence. A strong matching score (39.1% [Fig. 3C) was obtained between Orf4 and the N terminus of an ORF of yet unknown function already identified in both *E. coli* and

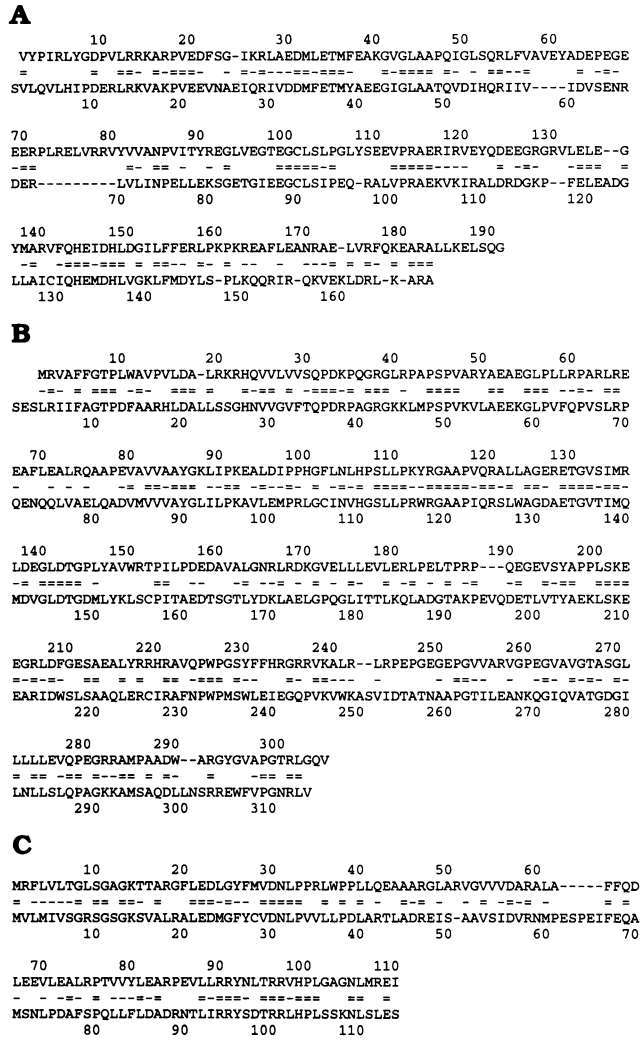

FIG. 3. Alignment of the three ORFs deduced from *T. thermophilus* DNA with the corresponding ones from *E. coli*. The amino acid sequences originating from *T. thermophilus* are shown on the top, and those from *E. coli* are shown on the bottom. PDFs (A), MTFs (B), and Orf4 (as defined in Fig. 2) (C) are compared in the respective panels. (C) Homologies were detected with the help of the FASTA program (22) by using the computer facilities of the Centre Interuniversitaire de Traitement de l'Information (CITI2) (5). Symbols: =, strictly identical; −, conservative replacement (taking into account the partition of amino acids into nine different classes defined as follows: A and G; S and T; D, E, N, and Q; K and R; I, L, M, and V; F, W, and Y; C; H; and P).

Klebsiella pneumoniae (9, 17). No significant homology could be found in the case of Orf1.

Nucleotide sequence of the *E. coli* region between the *fmt* and *trkA* genes. The nucleotide sequence of the region upstream from the *E. coli fms* gene (15) showed no identity with the corresponding region from *T. thermophilus*. To enlarge the comparison to the region downstream of the *E. coli fmt* gene, the nucleotide sequence of a 1,038-bp DNA fragment, corresponding to the region between *fmt* and *trkA*, was determined. Examination of this *E. coli* DNA sequence revealed two new ORFs, called *fmu* and *fmv* (Fig. 1B) and composed of 238 and 191 codons, respectively. They are transcribed clockwise, as are *fms*, *fmt*, and *trkA*. A putative rho-dependent transcription terminator is located between *fmt* and *fmu*. In turn, *fmu* and

fmv are likely to be transcribed together because the initiator ATG of the *fmv* ORF overlaps the TGA stop codon of *fmu*. Comparison of the amino acid sequences deduced from these two ORFs to the NBRF protein data bank (release 37) revealed no significant homology. Moreover, none of these two ORFs resembles the ORF flanking the *T. thermophilus fmt* gene. Apparently, although the PDF and MTF genes are linked in a similar manner on the DNA of the two bacteria, their genetic contexts are different.

Concluding remarks. In procaryotes, the removal of the N-terminal methionine from nascent proteins occurs through a mechanism involving the sequential action of two enzymes: PDF and methionine aminopeptidase. In this report, the *E. coli fms* gene which encodes PDF is shown to be required for cellular viability. This result is in agreement with the recent demonstration that an *E. coli* strain with a deletion of the whole *fms-fmt* locus could be transformed by a plasmid expressing MTF only if an inhibitor of the formylation reaction was added to the growth medium (13). The essential character of *fms* was expected, because deformylation is a prior requirement for the action of methionine aminopeptidase and because its gene is essential to *E. coli* cell growth (3). Alternative mechanisms sustaining the essential character of the *fms* gene may also be considered. First, deformylation of at least one protein, even if it is not processed by methionine aminopeptidase, may be required for the growth of the bacterium. Second, PDF might be necessary because it releases formate molecules from nascent polypeptides (1). Actually, in *E. coli*, formate is a substrate of the *purT* gene product in the formylation of 5′-phosphoribosyl-1-glycinamide to 5′-phosphoribosyl-*N*-formylglycinamide (21). To our knowledge, with PDF, the *purU* gene product is the only other candidate for a potential source of formate in *E. coli* (20).

The construction of a conditional-lethal *fms* strain allowed us to evidence by functional complementation the colocalization of the *fms* and *fmt* genes in *T. thermophilus*. The conservation of such a genetic linkage in both *E. coli* and *T. thermophilus* highlights the biological significance of the coexpression of MTF and PDF activities in procaryotic organisms. Since the processing of N-terminal methionines by methionine aminopeptidase is required, the occurrence of a formylation step at the initiation of translation implies that a deformylation step must be immediately ensured. In agreement with this, PDF activity is no longer required in an *E. coli* strain deprived of MTF (13). It can therefore be expected that a colocalization of the genes encoding PDF and MTF will be a constant trait of the procaryotic kingdom. In this context, the case of mitochondrial and chloroplastic MTF and PDF, which would be produced from genetic loci where transcriptional units are not likely to occur, deserves interest.

Nucleotide sequence accession numbers. The nucleotide sequence of the 2,292-bp DNA fragment, corresponding to the *fms-fmt* region from *T. thermophilus*, has been assigned EMBL data, library accession number X79087. The nucleotide sequence of the 1,038-bp DNA fragment, corresponding to the *E. coli* region between *fmt* and *trkA*, has been assigned EMBL data library accession number X77091.

We thank Y. Mechulam and N. G. Wallis for critical reading of the manuscript; J.-M. Guillon, who initiated the inactivation of the *fms* gene; and C. Lazennec for expert technical help.

REFERENCES

1. **Adams, J. M.** 1968. On the release of the formyl group from nascent protein. J. Mol. Biol. **33**:571–589.
2. **Bedwell, D., G. Davis, M. Gosink, L. Post, M. Nomura, H. Kestler, J. M. Zengel, and L. Lindahl.** 1985. Nucleotide sequence of the alpha ribosomal protein operon of *Escherichia coli*. Nucleic Acids Res. **13**:3891–3903.
3. **Chang, S.-Y. P., E. C. McGary, and S. Chang.** 1989. Methionine aminopeptidase gene of *Escherichia coli* is essential for cell growth. J. Bacteriol. **171**:4071–4072.
4. **Csonka, L. N., and A. J. Clark.** 1980. Construction of an Hfr strain useful for transferring *recA* mutations between *Escherichia coli* strains. J. Bacteriol. **143**:529–530.
5. **Dessen, P., C. Fondrat, C. Valencien, and C. Mugnier.** 1990. BISANCE: a French service for access to biomolecular sequence databases. Comput. Applic. Biosci. **6**:355–356.
6. **Guillon, J.-M., Y. Mechulam, J.-M. Schmitter, S. Blanquet, and G. Fayat.** 1992. Disruption of the gene for Met-tRNA$_f^{Met}$ formyltransferase severely impairs the growth of *Escherichia coli*. J. Bacteriol. **174**:4294–4301.
7. **Hamilton, C. M., M. Aldea, B. K. Washburn, P. Babitzke, and S. R. Kushner.** 1989. New method for generating deletions and gene replacements in *Escherichia coli*. J. Bacteriol. **171**:4617–4622.
8. **Hirel, P.-H., F. Lévêque, P. Mellot, F. Dardel, M. Panvert, Y. Mechulam, and G. Fayat.** 1988. Genetic engineering of methionyl-tRNA synthetase: *in vitro* regeneration of an active synthetase by proteolytic cleavage of a methionyl-tRNA synthetase-β-galactosidase chimeric protein. Biochimie (Paris) **70**:773–782.
9. **Jones, D. H., C. F. Franklin, and C. M. Thomas.** Unpublished data.
10. **Jongeneel, C. V., J. Bouvier, and A. Bairoch.** 1989. A unique signature identifies a family of zinc-dependent metallopeptidases. FEBS Lett. **242**:211–214.
11. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
12. **Livingston, D. M., and P. Leder.** 1969. Deformylation and protein synthesis. Biochemistry **8**:435–443.
13. **Mazel, D., S. Pochet, and P. Marlière.** 1994. Genetic characterization of polypeptide deformylase, a distinctive enzyme of eubacterial translation. EMBO J. **13**:914–923.
14. **Meinnel, T., and S. Blanquet.** 1993. Evidence that peptide deformylase and methionyl-tRNA$_f^{Met}$ formyltransferase are encoded within the same operon in *Escherichia coli*. J. Bacteriol. **175**:7737–7740.
15. **Meinnel, T., J.-M. Guillon, Y. Mechulam, and S. Blanquet.** 1993. The *Escherichia coli fmt* gene, encoding methionyl-transfer RNA$_f^{Met}$ formyltransferase, escapes metabolic control. J. Bacteriol. **175**:993–1000.
16. **Meinnel, T., Y. Mechulam, and S. Blanquet.** 1993. Methionine as translation start signal: a review of the enzymes of the pathway in *Escherichia coli*. Biochimie (Paris) **75**:1061–1075.
17. **Merrick, M. J., and J. R. Coppard.** 1989. Mutations in genes downstream of the *rpoN* gene (encoding σ^{54}) of *Klebsiella pneumoniae* affects expression from σ^{54}-dependent promoters. Mol. Microbiol. **3**:1765–1775.
18. **Miller, C. G., A. M. Kukral, J. L. Miller, and N. Rao Movva.** 1989. *pepM* is an essential gene in *Salmonella typhimurium*. J. Bacteriol. **171**:5215–5217.
19. **Miller, H. I., and D. I. Friedman.** 1980. An *E. coli* gene product required for λ site-specific recombination. Cell **20**:711–719.
20. **Nagy, P. L., G. M. McCorkle, and H. Zalkin.** 1993. *purU*, a source of formate for *purT*-dependent phosphoribosyl-*N*-formylglycinamide synthesis. J. Bacteriol. **175**:7066–7073.
21. **Nygaard, P., and J. M. Smith.** 1993. Evidence for a novel glycinamide ribonucleotide transformylase in *Escherichia coli*. J. Bacteriol. **175**:3591–3597.
22. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**: 2444–2448.
23. **Takeda, M., and R. E. Webster.** 1968. Protein chain initiation and deformylation in *B. subtilis* homogenates. Proc. Natl. Acad. Sci. USA **60**:1487–1494.
24. **Vallee, B. L., and D. S. Auld.** 1990. Zinc coordination, function, and structure of zinc enzymes and other proteins. Biochemistry **29**: 5647–5659.