Volume 17 Number 7 1989 Nucleic Acids Research

Occurrence and functional compatibility within Enterobacteriaceae of a tRNA species which inserts selenocysteine into protein

Johann Heider, Walfred Leinfelder and August Böck*

Lehrstuhl für Mikrobiologie der Universität München, Maria-Ward-Straße 1a, D-8000 München 19, FRG

Received February 14, 1989; Accepted March 3, 1989 EMBL accession no. X14255

ABSTRACT

The selC gene from E. coli codes for a tRNA species ($tRNA^{Ser}_{UCA}$) which is aminoacylated with L-serine and which co-translationally inserts selenocysteine into selenoproteins. By means of Southern hybridization it was demonstrated that this gene occurs in all enterobacteria tested. To assess whether the unique primary and secondary structural features of the E. coli selC gene product are conserved in that of other organisms, the selC homologue from Proteus vulgaris was cloned and sequenced. It was found that the Proteus selC gene differs from the E. coli counterpart in only six nucleotides, that it displays the same unique properties and that it is expressed and functions in E. coli. This indicates that the unique mechanism of selenocysteine incorporation is not restricted to E. coli but has been conserved as a uniform biochemical process.

INTRODUCTION

A biological role for the trace element selenium was first recognized in 1954 when J. Pinsent reported that gas formation by Escherichia (E.) coli depended upon the presence of selenite in the medium (1). It was revealed in subsequent biochemical studies that selenium is incorporated into two distinct polypeptides in E. coli which are constituent subunits of formate dehydrogenases: Formate dehydrogenase H (FDH_H) which is part of the formate-hydrogen-lyase complex possesses an 80 kD selenopolypeptide and formate dehydrogenase N (FDH_N), which delivers the electrons to nitrate reductase, possesses a 110 kD selenopolypeptide subunit (2, 3). It was shown for the 80 kD selenopolypeptide of FDH_H that selenium is incorporated co-translationally in the form of selenocysteine and that the in-corporation is directed by an in-frame UGA termination codon on the mRNA (4, 5).

The synthesis and incorporation of selenocysteine into these selenopolypeptides from E. coli requires the activity of the products of at least four genes, designated selA, selB, selC and selD (6). selA, selB and selD code for proteins whose function are still unknown (6), whereas selC codes for a tRNA species (7). This tRNA ($tRNA^{Ser}_{UCA}$) is aminoacylated with L-serine by seryl-tRNA ligase; it inserts selenocysteine into the protein after the L-serine, ester-bonded to $tRNA^{Ser}_{UCA}$, has been converted into a selenocysteine residue (8). Apart from its unique function, $tRNA^{Ser}_{UCA}$ (as delineated from the gene sequence which recently was confirmed by sequencing the gene product; D. Söll, personal communication) deviates in a number of unique features from the structure of those tRNA species of E. coli inserting one of the twenty classical amino acids. These are: (i) an 8-base pair amino-acyl acceptor stem; (ii) the anticodon directly matching the UGA termination codon; (iii) several deviations from positions hitherto considered as invariant within the D-stem region (7). It was hypothesized that these unique features might play a role in the decoding process of the UGA (determining selenocysteine insertion) or as a recognition signal for the enzyme(s) responsible for converting L-serine into selenocysteine (7).

Such functional essentiality should be correlated with a high evolutionary conservation of the structure. Therefore, we have searched for the occurrence of the selC homologue in other enterobacteria and determined the sequence of the selC homologous gene from Proteus (P.) vulgaris; an organism only distantly related to E. coli. We show that the organisms tested contain a gene homologous to the E. coli selC gene, that the Proteus selC gene sequence displays all the structural characteristics unique to the E. coli gene product and that the selC gene product from P. vulgaris is functional in E. coli.

MATERIAL AND METHODS

Bacterial strains and plasmids

The bacterial strains and plasmids used in this work are listed in Table 1. LB medium (9) was used for aerobic growth of bacteria; for anaerobic cultivation it was supplemented with

Table 1. Bacteria and plasmids used

Organism	Genotype	Source/Reference
E. coli		
MB07	F^- *proC23 trp-30 his-51 lac-28 rpsL101 selC* λ^+	(24)
MN1	like MB07 Δ(*srl-recA*)::Tn*10*	W. Leinfelder
MC4100	F^- Δ(*argF-lac*)U169 *flbB5301 araD139 ptsF25 deoC1 relA1 rpsL150* λ^-	(25)
FM420	like MC4100, *srl*::Δ Tn*10 recA56*	(4)
Salmonella typhimurium		
TL117	Δ (*proAB*)47	(26)
Proteus mirabilis	wild type	DSM 30115
Proteus vulgaris	wild type	DSM 2140
Klebsiella pneumoniae	wild type	laboratory strain
Klebsiella oxytoca M5a1	wild type	laboratory strain
Serratia marcescens	wild type	laboratory strain
Enterobacter cloacae	wild type	laboratory strain
Enterobacter aerogenes	wild type	laboratory strain

Plasmid	Relevant marker	
pACYC184	Cam^R Tet^R	(27)
pUC18	Ap^R *lacZ' lacI'*	(28)
pMN81[a]	Ap^R $selC_E$ *lacI'*	(7)
pPM2[b]	Ap^R $selC_P$ *lacI'*	
pPM3[b]	Ap^R $selC_P$ *lacI'*	
pPM5[b]	Ap^R $selC_P$ *lacI'*	

a) $selC_E$: *selC* gene from *E. coli*
b) $selC_P$: *selC* gene from *Proteus vulgaris*

0.4 % (w/v) glucose. When required, nitrate was added to 1 % (w/v) final concentration.

Recombinant DNA techniques

Standard recombinant DNA techniques were performed as described by Maniatis et al. (10). Size fractionation of re-

striction enzyme-generated DNA fragments was performed by electrophoresis in 0.8 % agarose gels; fragments were recovered from the gel either by electroelution employing a Biometra chamber, as recommended by the manufacturers, or by adsorption to glass powder (11). Transformation of enterobacteria was carried out following the RbCl procedure (12).

DNA sequencing

Sequencing of 5'- and 3'-end labelled DNA fragments was carried out according to (13). Polylinker sites bordering the DNA inserts to be sequenced, as well as the StyI site within the Proteus selC gene, were used for labelling.

Plate assay for formate dehydrogenase activity

The benzylviologen dye overlay technique was used in screening transformants for formate dehydrogenase activity (14).

In vivo labelling experiments with [^{75}Se]selenite

Labelling of cells with [^{75}Se]selenite (150 - 1,500 µCi/µmol) was performed under anaerobic conditions employing the technique of Cox et al. (2). Extracts of labelled cells were subjected to SDS polyacrylamide gel electrophoresis according to Laemmli et al. (15) and the labelled macromolecules were visualized by autoradiography.

Preparation and aminoacylation of tRNA

Bulk tRNA was prepared from the 100,000 g supernatant by phenol extraction and fractionated isopropanol precipitation (16). The purity of the preparation was checked by agarose gel electrophoresis. 100 µg of bulk tRNA was charged to completion with the respective amino acids in the reaction mixture given by Buckel et al. (17) using 50 µg protein of a 100,000 g supernatant.

RESULTS

Ubiquitous distribution of the selC gene in enterobacteria

The existence of the selC gene, which codes for a tRNA species inserting selenocysteine into proteins, was discovered in E. coli (7); its presence in other organisms has not been demonstrated up to now. To determine whether the pathway of selenoprotein synthesis outlined for E. coli is of more wide-

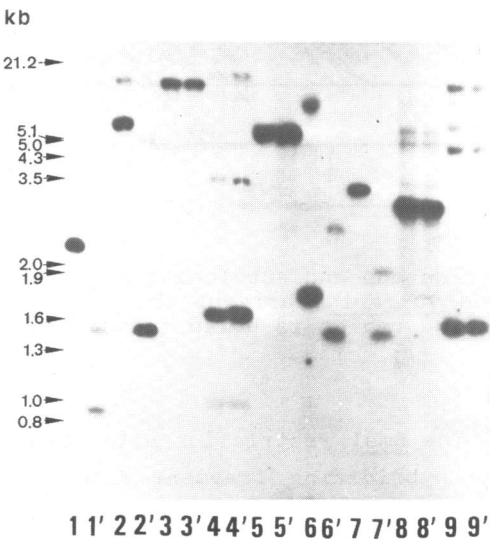

Fig. 1: Occurrence of a gene homologous to E. coli selC in enterobacteria. Chromosomal DNA from the respective organisms was fragmented with restriction endonucleases EcoRI plus SalI (lanes 1 to 9) or EcoRI, SalI plus SmaI (lanes 1' to 9'); the fragments were separated electrophoretically, transferred to nitrocellulose and hybridized with a radioactively labelled EcoRI/PstI fragment (450 bp) from plasmid pMN81 (7). The lanes contain DNA from: 1/1' E. coli MC4100; 2/2' Salmonella typhimurium TL117; 3/3' Proteus mirabilis DSM 30115; 4/4' Proteus vulgaris DSM 2140; 5/5' Serratia marcescens; 6/6' Klebsiella pneumoniae; 7/7' Klebsiella oxytoca M5a1; 8/8' Enterobacter cloacae; 9/9' Enterobacter aerogenes.

spread biological significance we have searched for the existence of a gene homologous to the E. coli selC gene in other enterobacteria. To this end, chromosomal DNA from several genera of this family was digested with restriction endonucleases EcoRI and SalI, which, in the case of E. coli, produce a 2.3 kb fragment carrying the selC gene (7). Aliquots of the digests were additionally incubated with SmaI to test whether the SmaI site present within the E. coli selC gene is conserved in the other organisms. Fig. 1 shows the hybridization pattern when a 450 bp fragment of plasmid pMN81 (7) is used as a probe. Intensive signals were obtained with all the organisms analysed which indicates that conserved sequences homologous to selC exist and are present as a single copy on

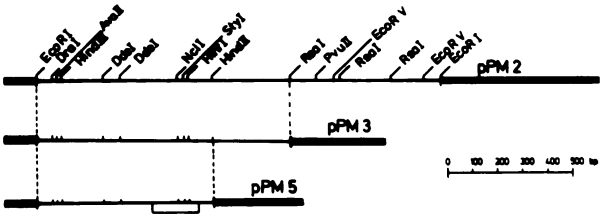

Fig. 2: Restriction map and subcloning strategy of the DNA fragment from Proteus vulgaris carrying the selC gene. The position of the selC encoding region is boxed in the map of plasmid pPM5.

the chromosome. The SmaI restriction site appears to be present only in the hybridizing fragment from Salmonella typhimurium (lane 2') and from the two Klebsiella species (lanes 6' and 7') which, taxonomically, are closer to E. coli than the other organisms tested (18).

Cloning and nucleotide sequence of the Proteus vulgaris selC gene

The P. vulgaris selC gene was chosen for a more detailed analysis since (i) this organism is only distantly related to E. coli (GC content of 39 % relative to 52 % for E. coli) (18) which increases the probability of variations in functionally non-essential sequence positions, and (ii) its selC gene is located on a conveniently short 1.6 kb DNA fragment (Fig. 1, lane 4). A fragment of this size was also detected in a Southern hybridization of P. vulgaris chromosomal DNA digested with endonuclease EcoRI alone (data not shown).

For cloning of the P. vulgaris selC gene, chromosomal DNA of this organism was digested with EcoRI and size-separated in agarose gels. Fragments of the appropriate size were extracted, ligated into the multilinker site of vector pUC18 and the ligation mixture was used for transformation of E. coli MN1 (selC). The transformants were screened for formate-dependent benzylviologen reduction with the aid of the plate dye overlay technique (14). Four transformants were obtained which had acquired plasmids with identical 1.6 kb EcoRI inserts. Upon retransformation of MN1 they restored to this strain formate

a

```
          -35                    -10
CTTGGGATTAAAAATAATTCAGTGTATAATCATCTCGTCATTTCCTTTTGACCCATTCAAAAAT
  **     *   *             ******   * *  *   *** * *  *  **  *
GATGTAGAAACTCAAGGAAGTAGCTATAAT--GCGCCCCGCCTCCATGTAGCA-ATCGAGGCGC

GGAAGATCATCGTCCCCGGTGGGGCGGCTGGACTTCAAATCCAGATGGGGCCGCCAGCGGTCCT
******** ***** ****** ********************** ******************
GGAAGATCGTCGTCTCCGGTGAGGCGGCTGGACTTCAAATCCAGTTGGGGCCGCCAGCGGTCCC

TGGCAGGTTCGACTCCTGTGATCTTCCGCCAATTAAGTTCAATACAACTCCACAAAACTCAAGT
 ********************************      **     * * * * ** * *  **
GGGCAGGTTCGACTCCTGTGATCTTCCGCCAAAATGCCTCTCCTGACGTTCTCCAACCCCTTGT

AATCCCTTATAATACCAGCCATTAATACACTTCCCAGT      P. vulgaris selC
*** *  *  *     *    **      **    *
AATACTGTCAATCCAGATATTTTCTACTGCTCAATATT      E. coli selC
```

b

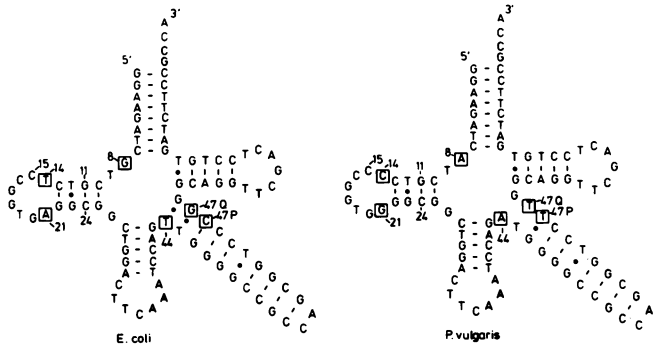

Fig. 3: Comparison of the selC genes from E. coli and Proteus vulgaris. (a) Alignment of the primary structures showing the non-coding strand in the 5' to 3' direction. The positions conserved in the two sequences are denoted by asterisks; the coding area and putative promoter motifs are underlined. (b) Putative secondary structures of the selC gene products; non-identical sequence positions are boxed. Numbers are according to the standard nomenclature (20).

dehydrogenase activity and the capacity for gas formation. The physical map of the insert of one of these plasmids, pPM2, is given in Fig. 2 together with the subcloning strategy which ultimately delivered plasmid pPM5 carrying a 0.8 kb insert still capable of complementing the selC lesion of MN1. Further subcloning revealed that the 366 bp DdeI-HindII fragment (see Fig. 2) cloned into plasmid pACYC184 expressed the complementing gene product, whereas restriction at the StyI site led to loss of complementation (data not shown).

The nucleotide sequence of the insert of plasmid pPM5 was

Table 2. Amino acid acceptance activity of bulk tRNA prepared from E. coli FM420 carrying plasmid pUC18, pPM5 or pMN81

Enzyme source[a]	amino acid	pmoles amino acid charged to 100 µg tRNA		
		pUC18	pPM5	pMN81
E. coli FM420	serine	63	106	180
	phenylalanine	61	40	59
	cysteine	20	17	n.d.[b]
	leucine	115	73	n.d.
	glycine	71	40	n.d.
	alanine	80	53	n.d.
	tryptophan	120	66	n.d.
	aspartic acid	72	45	n.d.
P. vulgaris	serine	59	104	172
	phenylalanine	46	33	55
	tryptophan	108	57	n.d.

a) 50 µg protein of a 100,000 g supernatant
b) n.d.: not determined

determined (Fig. 3a). The insert contains a stretch of DNA which is homologous to the E. coli selC gene sequence. There are only six nucleotide positions within the coding region which are different in the Proteus putative selC gene sequence when compared to that of E. coli; the sequence similarity breaks down to statistically insignificant levels upstream and downstream of the coding area. Fig. 3b illustrates that the P. vulgaris selC gene product can form a secondary structure identical to that of the E. coli selC gene product.

The P. vulgaris selC gene product functions in E. coli

The initial biochemical evidence which indicated that the P. vulgaris selC gene product is functional in E. coli came from the fact that the formate dehydrogenase lesion of mutant MN1 from E. coli was relieved by the cloned Proteus selC gene.

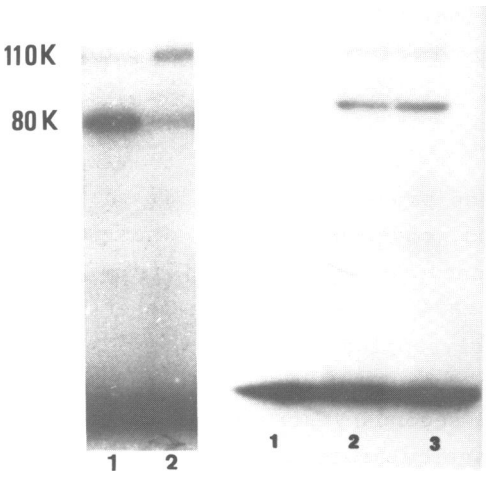

Fig. 4: Functional compatibility of the E. coli and Proteus vulgaris selC gene products. (a) [^{75}Se]selenium labelling of anaerobically grown P. vulgaris cells. Lane 1: growth in the absence of external electron acceptors; lane 2: growth in the presence of 1 % nitrate. (b) ^{75}Se incorporation into the selenopolypeptides of FDH_H and FDH_N by the selC mutant MN1 carrying vector pUC18 (lane 1), and by MN1 carrying the plasmids pPM5 (lane 2) and pMN81 (lane 3) which harbor the Proteus vulgaris and E. coli selC gene, respectively. Growth was in the absence of electron acceptor.

This functional compatibilty was analysed in more detail by testing whether the selC gene from this organism codes for a tRNA species which is aminoacylated by L-serine. Bulk tRNA was prepared from E. coli FM420 harboring plasmids pUC18, pPM5 or pMN81, and charged to completion with either one of several radioactive amino acids using 100,000 g extracts from E. coli or P. vulgaris (Table 2). The results show that the presence of the selC gene from E. coli or P. vulgaris on a multicopy plasmid specifically increased the acceptor activity for L-serine. Finally, it was analysed whether introduction of the Proteus selC gene into E. coli MN1 restored this strains' capacity for synthesis of the 80 kD and 110 kD selenoproteins. Fig. 4a shows that P. vulgaris, like E. coli, synthesises two selenopolypeptides which exhibit the same respective regulatory characteristics. Part (b) demonstrates that MN1 is pleiotropically defective in the formation of these selenoproteins and

that transformation with plasmids carrying the selC gene restored selenoprotein formation, irrespective of whether the gene was from E. coli or P. vulgaris.

DISCUSSION

The selC gene product, $tRNA^{Ser}_{UCA}$, is the key molecule in the biosynthesis and co-translational incorporation of selenocysteine. Selenocysteine is formed from an L-serine residue charged to this tRNA (7, 8), which inserts this unusual amino acid into the polypeptide chain at the position of an in-frame UGA codon of the message (4, 5). Therefore, $tRNA^{Ser}_{UCA}$ must possess primary and secondary structural features which ensure (i) that only the particular serine residue charged to it (and not to any cognate serine-specific tRNA) is converted to selenocysteine, and (ii) that the particular UGA codon of the message is differentiated from a normal UGA termination codon. The latter effect could be envisaged as taking place either via a direct interaction of some part of the tRNA molecule with the message or the ribosome or, alternatively, via some structural feature which alters the interaction pattern of this tRNA with translation factors (19).

Any of the unique structural features reported for the selC gene product (7) could be involved in these processes, particularly since they are conserved in the structure of the Proteus gene product: (i) By an A - U insertion the aminoacyl-acceptor stem of the selC tRNA from both organisms is extended to eight base pairs; (ii) both molecules contain an anticodon able to read the UGA codon; (iii) the Proteus tRNA displays the same deviations in invariant and semi-invariant nucleotides as the E. coli $tRNA^{Ser}_{UCA}$. These are in particular: the lack of a uridine in the position 8 (standard nomenclature) which is occupied by a G (E. coli) or A (P. vulgaris) and the presence of a purine-pyrimidine pair in positions 11/24 and of two pyrimidine residues in positions 14 and 15 (standard nomenclature).

The Proteus selC sequence differs from the E. coli one in six out of 95 nucleotides. Three base changes are located at the base of the extra arm; they are responsible for loss of the

SmaI restriction site which is present in the E. coli selC gene. The other three nucleotides, which are invariant in other tRNA species, involve an exchange of U_8 by a purine and U_{14} to C_{14} and A_{21} to G_{21} transitions. It is interesting to note that none of the latter interspecies sequence differences match the consensus situation (20), thus reinforcing the notion that these deviations from consensus are functionally essential. It appears that the alterations in positions 14 and 21 (E. coli versus P. vulgaris) represent compensatory base changes within an extended D stem. Such a long five (or even six) base pair stem region of $tRNA^{Ser}_{UCA}$ deviates from the three base pairs found in all other serine acceptor tRNA's from E. coli (21). A similar alternative tRNA structure has also been suggested for the mammalian opal suppressor tRNA by Hatfield and coworkers (22).

Comparison of the selC gene product with other serine accepting tRNA species strongly supports the contention that the end of the amino acid acceptor stem participates in recognition by seryl-tRNA ligase (23). Identification of the biochemical function of other parts of $tRNA^{Ser}_{UCA}$, especially of the unique structural features, must await the synthesis and analysis of mutant molecules.

The results reported demonstrate that the unique mechanism of selenocysteine incorporation is not restricted to E. coli but occurs in other organisms. In view of recent results which indicate that the O-phoshorylseryl-tRNA from mammals actually inserts selenocysteine (29), it appears that the pathway of selenocysteine synthesis from a serine residue charged to tRNA, and the co-translational mode of incorporation, is phylogenetically old and has been conserved as a uniform biochemical process in different cell lineages.

ACKNOWLEDGMENT

We thank Eva Zehelein for expert technical assistance, Gary Sawers for many suggestions and for reading the manuscript und Dieter Söll and Dolph Hatfield for communicating unpublished results. This work was supported by grants from the Bundesministerium für Forschung und Technologie and the Fonds der Chemischen Industrie.

*To whom correspondence should be addressed

REFERENCES

1. Pinsent,J. (1954) Biochem. J. 57, 10-16.
2. Cox,J.C., Edwards,E.S. and DeMoss,J.A. (1981) J. Bacteriol. 145, 1317-1324.
3. Pecher,A., Zinoni,F. and Böck,A. (1985) Arch. Microbiol.36, 131-136.
4. Zinoni, F., Birkmann,A., Stadtman,T.C. and Böck,A. (1986) Proc. Natl. Acad. Sci. USA 83, 4650-4654.
5. Zinoni,F., Birkmann,A., Leinfelder,W. and Böck,A. (1987) Proc. Natl. Acad. Sci. USA 84, 3156-3160.
6. Leinfelder, W., Forchhammer, K., Zinoni, F., Sawers, G., Mandrand-Berthelot,M.A. and Böck,A. (1988a) J. Bacteriol. 170, 540-546.
7. Leinfelder,W., Zehelein,E. and Böck, A. (1988b) Nature 331, 723-725.
8. Leinfelder,W., Stadtman,T.C. and Böck, A. (1989) J. Biol. Chem. (in press).
9. Miller,J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N.Y.
10. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N.Y.
11. Vogelstein,B. and Gillespie,D. (1979) Proc. Natl. Acad.Sci. USA 76, 615-619.
12. Hanahan,D. (1985) In DNA cloning, Glover, D.M. Ed., Vol. I, pp. 109-175, IRL Press, Oxford.
13. Maxam,A.M. and Gilbert,W. (1980) Methods Enzymol. 65, 499-561.
14. Mandrand-Berthelot,M.A., Wee,M.J.K. and Haddock,B.A. (1978) FEMS Microbiol. Lett. 4, 37-40.
15. Laemmli,U.K. (1970) Nature 227, 680-685.
16. Zubay,G. (1962) J. Mol. Biol. 4, 347-356.
17. Buckel, P., Piepersberg, W. and Böck,A. (1976) Mol. Gen. Genet. 149, 51-61.
18. Brenner,D.J. (1984) In Bergey's Manual of Systematic Bacteriology, Krieg,N.R. Ed., Vol. I, pp. 408-516, Williams & Wilkins, Baltimore.
19. Böck,A. and Stadtman,T.C. (1988) Biofactors 1, 245-250.
20. Sprinzl,M., Hartmann,T., Meissner,F., Moll,J. and Vorderwühlbecke,T. (1987) Nucleic Acids Res. 15, r53-r188.
21. Grosjean,H., Nicoghosian,K., Haumout,E., Söll,D. and Cedergren, R. (1985) Nucleic Acids Res. 13, 5697-5706.
22. Hatfield,D., Diamond,A. and Dudock, B. (1982) Proc. Natl. Acad. Sci. USA 79, 6215-6219.
23. Rogers,M.J. and Söll,D. (1988) Proc. Natl. Acad. Sci. USA 85, 6627-6631.
24. Haddock,B.A. and Mandrand-Berthelot,M.A. (1982) Bioch. Soc. Trans. 10, 478-480.
25. Casadaban,M.J. and Cohen,S.N. (1979) Proc. Natl. Acad. Sci. USA 76, 4530-4533.
26. Csonka,L.N. (1981) Mol. Gen. Genet. 182, 82-86.
27. Chang,A.C.Y. and Cohen, S.N. (1978) J. Bacteriol. 143, 971-980.
28. Norrander,J., Kempe,T. and Messing,J. (1983) Gene 26, 101-106.
29. Lee,B.J., Worland,P.J., Davis,J.N., Stadtman,T.C. and Hatfield,D.L. (1989) J. Biol. Chem. (in press).