Molecular cloning and sequencing of *pheU*, a gene for *Escherichia coli* tRNA^{Phe}

Ira Schwartz[1], Robin-Ann Klotsky[1], Dirk Elseviers[2], Patricia J.Gallagher[2], Manuel Krauskopf[3]*, M.A.Q.Siddiqui[3], James F.H.Wong[4] and Bruce A.Roe[4]

[1]Departments of Biochemistry and [2]Microbiology, New York Medical College, Valhalla, NY 10595, [3]Roche Institute of Molecular Biology, Nutley, NJ 07110, and [4]Department of Chemistry, University of Oklahoma, Norman, OK 73019, USA

Received 6 April 1983; Revised and Accepted 1 June 1983

ABSTRACT

A recombinant plasmid (designated pID2) carrying the E. coli gene for tRNA^{Phe} has been isolated from a plasmid bank constructed by the ligation of a total EcoRI digest of E. coli K12 DNA into the EcoRI site of pACYC184 DNA. The plasmid was selected by virtue of its ability to complement a temperature-sensitive lesion in the gene (PheS) for the α-subunit of phenylalanyl-tRNA synthetase. Crude tRNA isolated from such transformants exhibited elevated levels of phenylalanine acceptor activity. The tRNA^{Phe} gene has been localized within the first 300 base pairs of a 3.6 kb SalI fragment of pID2. The sequence of the gene and its flanking regions is presented.

INTRODUCTION

In recent years, the location and organization of many transfer RNA genes in E. coli have been elucidated (1,2). The genes are distributed throughout the genome (3) and occur either as multiple copies of identical genes (4), in units containing genes for several different tRNAs (5,6), as part of ribosomal RNA operons (7,8) and in transcription units with mRNA (9,10). To date, however, the expression of only a few E. coli tRNA genes has been studied in detail. These include the tRNA^{Tyr} genes (9,11) and those which are co-transcribed with ribosomal RNA (8,12).

In an earlier report the isolation of two plasmids, pID1 and pID2, which could confer temperature resistance to E. coli NP37, a temperature-sensitive mutant for pheS, was described (13). A plasmid bank containing EcoRI fragments representing the entire E. coli genome was used to transform E. coli NP37 and colonies capable of growth at 42° (the non-permissive temperature) were selected. Transformants harboring pID1 overproducted phenylalanyl-tRNA synthetase and pID1 DNA was shown to contain the genes thrS, infC, pheS and pheT (these genes code for threonyl-tRNA synthetase, initiation factor 3, and the α and β subunits of phenylalanyl-tRNA synthetase, respectively). In extracts of pID2-

transformed cells, however, no overproduction of any cytoplasmic protein was observed.

In our previous report we stated that the levels of $tRNA^{Phe}$ were elevated in NP37(pID2). Here we present evidence that the gene for $tRNA^{Phe}$ is contained on pID2 DNA. The sequence of the gene and its 5' and 3' flanking regions has been determined.

MATERIALS AND METHODS

Bacterial Strains. NP37, a pheS temperature-sensitive mutant (14), was obtained from the E. coli Genetic Stock Center.

Media. Bacteria were routinely grown on L medium (13) supplemented with 5 μg/mL of tetracycline or 30 μg/mL chloramphenicol.

Nucleic Acids. Large scale purification of plasmid DNA was performed as described (13). Rapid extraction of plasmid DNA was carried out according to Birnboim and Doly (15) from overnight cultures grown in LTc. Crude tRNA was isolated as described by Zubay (16).

Southern Transfer and Hybridization. DNA was subjected to digestion by restriction endonucleases for 16 hours and electrophoresed on agarose gels as previously described (13). The restriction map was constructed on the basis of double digestion patterns or according to the procedure of Smith and Birnstiel (17).

DNA was transferred from the gels to nitrocellulose filters and hybridized with ^{32}P-labeled $tRNA^{Phe}$ essentially as described by Southern (18). Purified E. coli tRNA (Boehringer) was dephosphorylated by treatment with bacterial alkaline phosphatase (Bethesda Research Laboratories) for one hour at 65°C in 30 mM Tris-HCl (pH 8), 0.1% sodium dodecyl sulphate. Dephosphorylated $tRNA^{Phe}$ (55 μg) was labeled at the 5'-end by incubation for 40 minutes at 37° in a mixture containing 1 mCi γ-^{32}P-ATP(3000 Ci/mmol), 20 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 15 mM β-mercaptoethanol and 0.6 units of T_4 polynucleotide kinase. The resultant $tRNA^{Phe}$ had a specific activity of approximately 1 x 10^8 cpm/ μg. Hybridization was carried out overnight at 37° and the filter was washed with 2 x SSPE [0.18 M NaCl, 10 mM $NaPO_4$, 1.0 mM Na_2EDTA (pH 7)] at 45° for one hour and then four times with 0.2 x SSPE, 0.3% SDS at 65°C for a total of one hour. Autoradiography was carried out at -70°C with Kodak X-AR film and Dupont Cronex Lightning Plus intensifer screens.

Aminoacylation of tRNA. The standard reaction mixture for aminoacylation of tRNA consisted of 100 mM Na.Hepes (pH 7.5), 10 mM KCl, 40 mM $MgCl_2$, 1.0 mM

dithiothreitol, 5 mM ATP, 10-20 μM of radioactive amino acid and saturating amounts of a crude preparation of E. coli aminoacyl-tRNA synthetases [prepared according to Muench and Berg (19)]. tRNA was present in limiting amounts and the charging levels were determined from the linear portions of a tRNA titration curve.

Cloning in pBR322 and Bacteriophage M13. A 3.6 kb SalI fragment from pID2, containing the gene for $tRNA^{Phe}$, was inserted into the SalI site of pBR322. This plasmid is designated pRK3. This fragment was excised from pRK3 and subcloned into bacteriophage M13mp9.

DNA Sequencing. All DNA sequences were obtained from M13 subclones of the 3.6 kb SalI fragment by the dideoxynucleotide chain termination method (20).

RESULTS

Restriction Endonuclease Mapping of pID2 DNA. The construction of pID2, a hybrid plasmid containing an EcoRI fragment of the E. coli chromosome, was described earlier (13). pID2 DNA was subjected to digestion with a number of restriction endonucleases. The resulting restriction map is presented in Figure 1. Comparison of the maps of pID1 and pID2 indicates that the 9.5 kb insert in pID2 is unrelated to the 22 kb insert previously identified in pID1 (13).

Amino Acid Acceptor Activities of tRNA Isolated from Transformed E. coli NP37. E. coli NP37 is a mutant with a temperature-sensitive lesion in pheS, the gene coding for the α-subunit of phenylalanyl-tRNA synthetase. Since pID2 did not contain pheS it was assumed that the plasmid might contain the gene for another participant in the phenylalanine charging reaction. In order to test this possibility bulk tRNA was isolated from NP37(pACYC184) and NP37(pID2) and charged with phenylalanine as described in Methods. The results shown in Table 1 indicate that the phenylalanine acceptor activities of tRNA from NP37(pID2) and NP37(pRK3) are elevated five -and eight-fold, respectively.

Hybridization Analysis. In order to directly demonstrate that pID2 contained the gene for $tRNA^{Phe}$ pID2 DNA was subjected to hybridization analysis using $^{32}P\text{-}tRNA^{Phe}$ as probe. The results are shown in Figure 2. Digestion of pID2 DNA with AvaI, PstI or SalI followed by separation of the resulting fragments by gel electrophoresis and blotting to nitrocellulose in each case resulted in the hybridization of one restriction fragment to the $^{32}P\text{-}tRNA^{Phe}$ probe. The sizes of the hybridizing fragments were 6.6 kb, 7.0 kb and 3.6 kb for AvaI, PstI, SalI, respectively. Weak hybridization to the 5.0 kb AvaI fragment was also observed. This is probably due to the presence of a

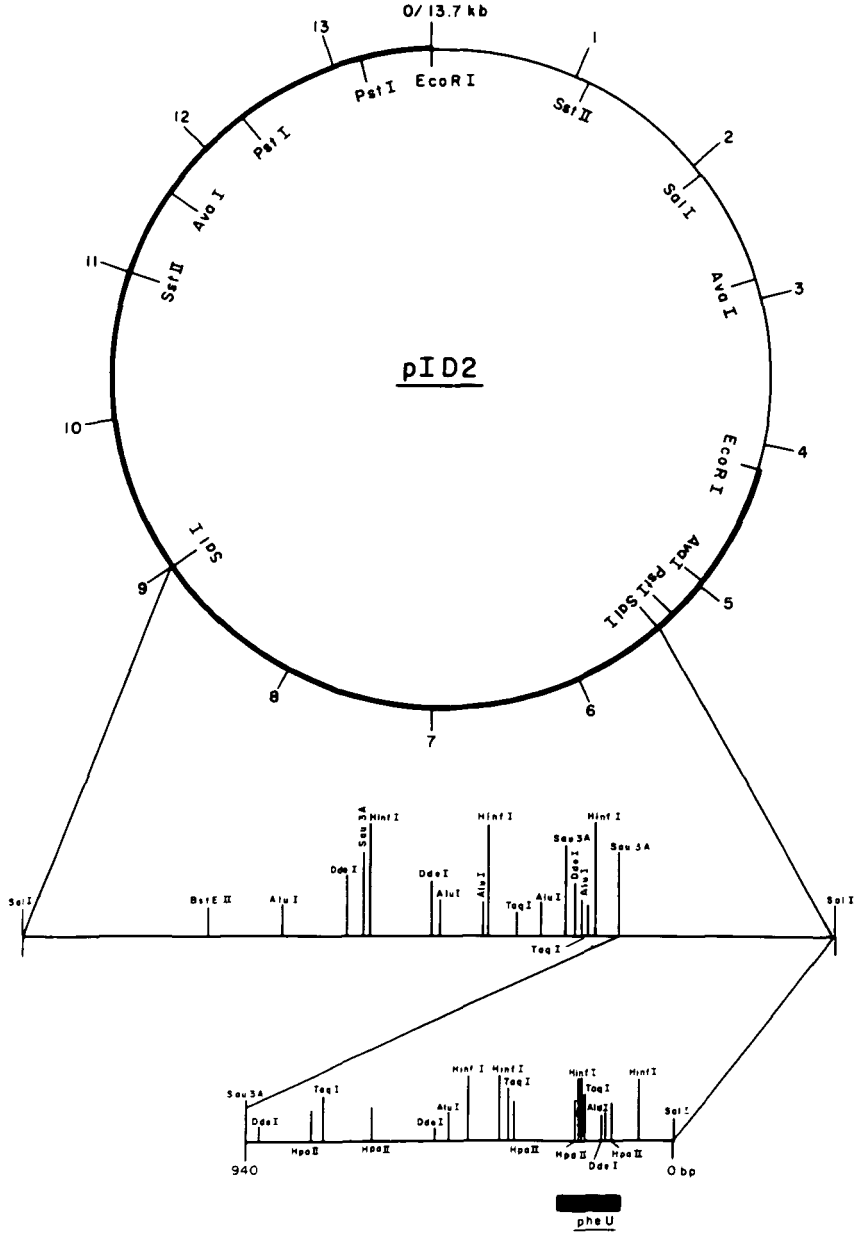

Figure 1. Restriction Map of pID2 DNA. The thick bar bounded by EcoRI sites is an insert of E. coli DNA. For clarity a more detailed restriction map of the 3.6 kb SalI fragment and a 940 b.p. SalI - Sau3A fragment are shown in a linear form. The black rectangle denotes the region of the SalI - Sau3A fragment which contains pheU.

Table 1 Phenylalanine acceptor activities of tRNA from plasmid-transformed E. coli NP37

Plasmid	Acceptor Activity (pmol/A260)	Overproduction (%)
pACYC184	36.2	-
pID2	180.6	499
pRK3	281.1	780

minor contaminating tRNA species in the $tRNA^{Phe}$ preparation used as the probe.

Since the sequence of $tRNA^{Phe}$ is known (21) it was possible to predict the restriction map for the coding portion of its gene. Detailed restriction analysis of the 3.6 kb SalI fragment yielded a map which was

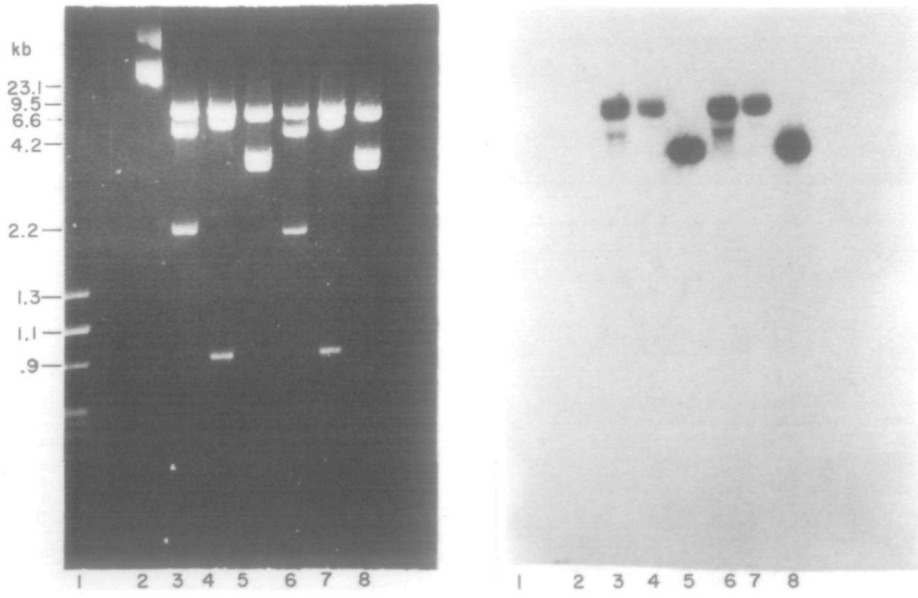

Figure 2. Hybridization of ^{32}P-$tRNA^{Phe}$ to restricted pID2 DNA. pID2 DNA was digested with the indicated enzymes, the fragments separated by 1% agarose gel electrophoresis, transferred to nitrocellulose and probed with ^{32}P-$tRNA^{Phe}$ as described in Methods. Left, photograph of ethidium bromide-stained gel; right, autoradiograph of filter after transfer and hybridization. Lane 1: Øx174 DNA digested with HaeIII; Lanes 2-8: pID2 DNA, undigested (lane 2); digested with AvaI (lanes 3 and 6); digested with PstI (lanes 4 and 7); digested with SalI (lanes 5 and 8).

```
-181
 GTCGACACTG CACATCGGCT GGCGGAAGAT ATCCTGCGCT TCGCCGCGTG GCCGCCAGGT
 CAGCTGTGAC GTGTAGCCGA CCGCCTTCTA TAGGACGCGA AGCGGCGCAC CGGCGGTCCA

-121
 TGGTGCATTG ACCTGACATA AACACAGAAA AGAAGCGATT TGCCGCAATC TTAAGCAGTT
 ACCACGTAAC TGGACTGTAT TTGTGTCTTT TCTTCGCTAA ACGGCGTTAG AATTCGTCAA

-61                        ********                    **** ***
 GAATCGCTTT TACTGACATT AGGTTGACGA GATGTGCAGA TTACGGTTTA ATGCGCCCGT
 CTTAGCGAAA ATGACTGTAA TCCAACTGCT CTACACGTCT AATGCCAAAT TACGCGGGCA

  +1
 TGCCCCGATA GCTCAGTCGG TAGAGCAGGG GATTGAAAAT CCCCGTGTCC TTGGTTCGAT
 ACGGGCCTAT CGAGTCAGCC ATCTCGTCCC CTAACTTTTA GGGGCACAGG AACCAAGCTA

+60
 TCCGAGTCCG GGCACCAAAT TACGCGGGGC AACCGGGGCC TGATCGAGTC AGCCATCTGG
 AGGCTCAGGC CCGTGGTTTA ATGCGCCCCG TTGGCCCCGG ACTAGCTCAG TCGGTAGACC

+120
 CCCCTAACTT TT
 GGGGATTGAA AA
```

Figure 3. Sequence of the $tRNA^{Phe}$ gene and its flanking regions. +1 is the beginning of the mature $tRNA^{Phe}$ sequence. This sequence is enclosed in a box. The sequences indicated by asterisks correspond to the Pribnow box and -35 regions. The overline denotes the inverted repeat sequence. As per convention, the upper line is written in the 5' to 3' direction.

consistent with the prediction and allowed the $tRNA^{Phe}$ gene to be localized within the first 300 b.p. from one end of the 3.6 kb fragment (Figure 1).

Sequence of the $tRNA^{Phe}$ Gene. The DNA sequence of each end of the 3.6 kb SalI fragment was determined. The sequence corresponding to the gene for $tRNA^{Phe}$ and its flanking regions was obtained by analysis of six different overlapping M13 clones. This sequence is presented in Figure 3.

DISCUSSION

The data presented here indicate that overproduction of $tRNA^{Phe}$ resulted in the suppression of a temperature-sensitive lesion in phe-tRNA synthetase. This is not surprising if it is assumed that the temperature-sensitivity of NP37 is the result of its inability to produce sufficient levels of phe-tRNA. It has been shown previously (22) that overproduction of $tRNA^{Gln}_1$ was responsible for the suppression of a temperature-sensitive glutaminyl-tRNA synthetase mutation. In that case the overproduction was the result of a gene duplication. In the present study this has been accomplished by cloning the tRNA gene in a multicopy plasmid.

The precise molecular defect in the phe-tRNA synthetase from NP37 which results in its temperature-sensitivity is not known. The lesion is in the α-subunit of the enzyme (23). It is tempting to speculate that the mutant enzyme has an altered K_M for $tRNA^{Phe}$ and that this defect can be overcome by overproduction of the substrate. This possibility is currently being tested. Alternatively, it is possible that the elevated levels of $tRNA^{Phe}$ protect the synthetase against thermal inactivation at the non-permissive temperature. If the enzyme does indeed have an altered K_M it would suggest that at least a portion of the $tRNA^{Phe}$ binding site is localized on the α-subunit.

Crude tRNA from NP37(pID2) was tested for overproduction of other tRNAs as measured by amino acid acceptor activities. All amino acids except for isoleucine and tryptophan were tested and, except for phenylalanine, the acceptor activities were identical to wild-type tRNA (data not shown). While pID2 may contain genes for other tRNAs which for some reason are not overproduced in a manner similar to $tRNA^{Phe}$, this possibility seems highly unlikely.

NP37(pRK3) tRNA has a phenylalanine acceptor activity of 281 pmol/A_{260} which represents an eight-fold enrichment over that found in wild-type $tRNA^{Phe}$. This should make NP37(pRK3) the strain of choice for the purification of $tRNA^{Phe}$. We are currently constructing other plasmids which will hopefully result in even greater production of $tRNA^{Phe}$.

The sequence presented here contains regions which are characteristic of RNA polymerase binding sites. 9-15 b.p. upstream of the mature tRNA is a TTTAATG sequence which differs at only one position from the consensus sequence, TATAATG, expected at a Pribnow box (24). At -35 is a GGTTGACG sequence which also has strong homology to the proposed consensus sequence. There are no other sequences within the 181 nucleotides upstream of the mature $tRNA^{Phe}$ sequence which correspond to those expected of a promoter. This suggests that transcription begins only three nucleotides before the mature $tRNA^{Phe}$. This is consistent with the sequence of a $tRNA^{Phe}$ precursor isolated from an RNase P mutant which was shown to be pppGUU-$tRNA^{Phe}$ (25).

Several other aspects of the sequence warrant mention. The sequence at -35 is preceeded by a region which is very AT-rich. There are eleven AT base pairs from positions -51 to -38. This is true to some extent for all promoters (24) but is an especially prominent feature of other rRNA and tRNA promoters (see below). It has been suggested that such AT-rich

	-40 -30 -20 -10 +1
pheU	TTAGGTTGACGAGATGTGCAGATTACGGTTTAATGCGCCCG
leuV	ACTATTGACGAAAAGCTGAAAACCACTAGAATGCGCCTCCG
supB-E	GAGGTTGACGCTGCAAGGCTCTATACGCATAATGCGCCCCG
thrU	TTAGTTGCATGAACTCGCATGTCTCCATAGAATGCGCGCTA
tryT	AACACTTTACAGCGGCGCGTCATTTGATATGATGCGCCCCG
rrnA	CCTCTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCA
rrnD	ATACTTGTGCAAAAAATTGGGATCCCTATAATGCGCCTCCG
rrnE	TCTATTGCGGCCTGCGGAGAACTCCCTATAATGCGCCTCCA
rrnX	CCGCTTGTCTTCCTGAGCCGACTCCCTATAATGCGCCTCCA

Figure 4. Sequences of promoter regions of tRNA and rRNA genes in E. coli. Position +1 is the transcription start site. Sequence homologies in the -10 and -35 regions are underlined. References to sequences are pheU (this work), leuV (28), supB-E (6), thrU (29), tyrT (30), rrnA and rrnE (26), rrnD and rrnX (31).

regions may melt readily and thus facilitate the binding of RNA polymerase (26). In addition, the region between the Pribnow box and the transcription start site is rich in GC base pairs. This is also a characteristic of other rRNA and tRNA genes (see below) and may be related to the control of the transcription of these genes by ppGpp (26,27).

A number of other tRNA genes have been sequenced and it is instructive to compare the likely transcription control elements for sequence homologies. The 5' flanking sequences for leuV (28), supB-E (6), thrU (29), tyrT (30) and pheU (this work) are presented in Figure 4. In addition, the corresponding regions of the ribosomal RNA genes rrnA (26), rrnD (31), rrnE (26) and rrnX (31) are shown. Several striking instances of homology are apparent. A Pribnow box consensus sequence for these nine genes is TATAATG. The AATG is found in all cases except tyrT. This hepatanucleotide is identical to the consensus proposed by Rosenberg and Court (24) but what is particularly striking is the invariance in the genes considered here. There is also homology, although less strict, in the -35 region. The canonical TTG sequence (24) is found in all the genes except tyrT. This homology can be extended to TTGACG for pheU, leuV and supB-E. Indeed, pheU and supB-E have an identical GGTTGACG sequence in this

region. It is interesting to note that thrU and tyrT show very little homology to the other three tRNA genes. Whether the TTGACG sequence has any special function relating to the transcription of tRNA genes must await the elucidation of additional tRNA gene sequences. Finally, as indicated earlier, there is a high density of GC base pairs between -10 and the transcription start site. In particular there is an almost invariant CGCC sequence immediately downstream of the Pribnow box.

The sequence downstream of the mature tRNA contains a ten base pair inverted repeat which is separated by ten base pairs. This 30 base pair sequence (94-123) should be capable of forming a stable stem-loop structure since the stem would include six GC base pairs. This is followed by a run of four T residues, the first of which occurs 20 nucleotides from the center of dyad symmetry. These are all characteristic features of a transcription termination site (24). A similar transcription termination site was observed for the supB-E operon and in that case it was shown to function as a rho-independent terminator in vitro (32). In contrast, tyrT has a rho-dependent termination site (11) whose features are unlike those identified here. In vitro transcription studies will confirm whether this region in pheU does indeed function in rho-independent termination. It thus appears that the pheU operon contains a single tRNA gene since both promoter and terminator sequences surround a single $tRNA^{Phe}$ sequence.

After this work was completed and being prepared for publication a report appeared which described the cloning of pheU by a strategy similar to the one presented here (33). In that case it was also shown that pheU carried on pBR322 caused the overproduction of $tRNA^{Phe}$ and pheU was identified by hybridization analysis. Caillet, et al reported that $tRNA^{Phe}$ hybridizes to a 5.6 kb PstI fragment while in our case hybridization was to a 7.0 kb PstI fragment. The reason for this discrepancy is not clear and one cannot yet rule out the possibility that there are two genes for $tRNA^{Phe}$. In any event, the present communication provides further information in terms of the sequence of this gene and its control regions. These studies are being extended to an investigation of the regulatory mechanisms involved in the expression of pheU and of the molecular basis of the suppression of a thermosensitive pheS by overproduction of $tRNA^{Phe}$.

ACKNOWLEDGEMENTS

We thank Drs. M. Fournier, F. Neidhardt and D. Soll for helpful discussions. The participation of Leonard Scinto in the early stages of

this work is acknowledged.

This work was supported by National Institutes of Health Grants GM24576(to D.E.), GM29265(to I.S.) GM30399 and GM30400(to B.A.R.). D.E. is a Career Scientist awardee of the Irma T. Hirschl Charitable Trust. I.S. is a Sinsheimer Scholar awardee of the Alexandrine and Alexander L. Sinsheimer Foundation.

*Permanent Address: Inst. de Bioquimica, Fac. de Ciencias, Univ. Austral de Chile, Valdivia, Chile.

REFERENCES

1. Abelson, J. (1980) in "Transfer RNA: Biological Aspects," Söll, D., Abelson, J. and Schimmel, P.R., eds., pp. 211-219, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
2. Daniel, V. (1981) Crit. Rev. Biochem. 9, 253-292.
3. Cheung, A.Y.P. (1980) in "Transfer RNA: Biological Aspects," Söll, D., Abelson, J. and Schimmel, P.R. eds, pp. 563-564, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
4. Russell, R.L., Abelson, J., Landy, A., Gefter, M.L., Brenner, S. and Smith, J.D. (1970) J. Mol. Biol. 47, 1-13.
5. Squires, C., Konrad, B., Kirschbaum, J. and Carbon, J. (1973) Proc. Nat. Acad. Sci. 70, 438-441.
6. Nakajima, N., Ozeki, H. and Shimura, Y. (1981) Cell 23, 239-249.
7. Lund, E., Dahlberg, J.E., Lindhal, L., Jaskunas, S.R., Dennis, P.P. and Nomura, M. (1976) Cell 7, 165-177.
8. Morgan, E.A., Ikemura, T. and Nomura, M. (1977) Proc. Nat. Acad. Sci. 74, 2710-2714.
9. Hudson, L., Rossi, J. and Landy, A. (1981) Nature 294, 422-427.
10. Miyajima, A., Shibuya, M., Kuchino, Y. and Kaziro, Y. (1981) Mol. Gen. Genet. 183, 13-19.
11. Rossi, J., Egan, J., Hudson, K. and Landy, A. (1981) Cell 26, 305-314.
12. Ikemura, T. and Nomura, M. (1977) Cell 11, 779-793.
13. Elseviers, D., Gallagher, P., Hoffman, A., Weinberg, B. and Schwartz, I (1980) J. Bact. 152, 357-362.
14. Eidlic, L. and Neidhardt, F.C. (1965) J. Bact. 89, 706-711.
15. Birnboim, H.C. and Doly, J. (1979) Nucl. Acid Res. 7, 1513-1523.
16. Zubay, G. (1962) J. Mol Biol. 4, 347-356.
17. Smith, H.O. and Birnstiel, M.L. (1976) Nucl. Acid Res. 3, 2387-2398.
18. Southern, E. (1975) J. Mol. Biol. 98, 503-517.
19. Muench, K. and Berg, P. (1966): Procedures in Nucleic Acid Res. V. 1, Cantoni, G.L. and Davies, D.R., eds., pp. 375-383, Harper and Row, New York.
20. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161-178.
21. Barrell, B.G. and Sanger, F. (1969) FEBS Lett. 3, 275-278.
22. Morgan, S., Korner, A., Low, K.B. and Söll, D. (1977) J. Mol. Biol. 117, 1013-1031.
23. Comer, M.M. and Böck, A. (1976) J. Bact. 127, 923-933.
24. Rosenberg, M. and Court, D. (1979) Ann. Rev. Genet. 13, 319-353.
25. Vogeli, G., Grosjean, H. and Söll, D. (1975) Proc. Nat. Acad. Sci. 72, 4790-4794.
26. deBoer, H.A., Gilbert, S.F. and Nomura, M. (1979) Cell 17, 201-209.

27. Travers, A.A. (1980) J. Bact. 141, 973-976.
28. Duester, G., Campen, R.K. and Holmes, W.M. (1981) Nucl. Acids Res. 9, 2121-2139.
29. An, G. and Friesen, J. (1980) Gene 12, 33-39.
30. Sekiya, T., Takeya, T., Contreras, R., Kupper, H. Khorana, H.G. and Landy, A. (1976) in "RNA Polymerase," Losick, R. and Chamberlain, M., eds., pp. 455-472, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
31. Young, R.A. and Steitz, J. A. (1979) Cell 17, 225-234.
32. Nakajima, N., Ozeki, H. and Shimura, Y. (1982) J. Biol. Chem. 257, 11113-11120.
33. Caillet, J., Plumbridge, J., Springer, M., Vacher, J., Delamarche, C., Buckingham, R. and Grunberg-Manago, M. (1983) Nucl. Acids. Res. 11, 727-736.