Volume 8 Number 5 1980 Nucleic Acids Research

A phenylalanine tRNA gene from Neurospora crassa: conservation of secondary structure involving an intervening sequence

Eric Selker and Charles Yanofsky

Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA

Received 27 November 1979

ABSTRACT

We have isolated and sequenced a $tRNA^{Phe}$ gene from Neurospora crassa. Hybridization analyses suggest that $tRNA^{Phe}$ is the only tRNA encoded on the cloned 5 kb DNA fragment. The $tRNA^{Phe}$ gene contains an intervening sequence 16 nucleotides in length located one nucleotide 3' to the anticodon position. The $tRNA^{Phe}$ coding region of Neurospora and yeast are 91% conserved, whereas their intervening sequences are only 50% identical[1]. The pattern of sequence conservation is consistent with a proposed secondary structure for the tRNA precursor in which the anticodon is base paired with the middle of the intervening sequence and the splice points are located in adjacent single-stranded loops. The DNA sequence following the $tRNA^{Phe}$ coding region is similar to sequences following other genes transcribed by RNA polymerase III in that it is AT-rich and includes a tract of A residues in the coding strand[1-3]. In contrast, the sequence preceding the Neurospora $tRNA^{Phe}$ coding region does not resemble sequences preceding other sequenced tRNA genes[1,2].

INTRODUCTION

Some, but not all, eukaryotic tRNA genes contain intervening sequences. Goodman et al.[2] sequenced several yeast $tRNA^{Tyr}$ genes and found 14 nucleotides adjacent to the anticodon region which are not represented in the mature tRNAs. Similarly, Valenzuela et al.[1] discovered intervening sequences 18 or 19 nucleotides long in yeast $tRNA^{Phe}$ genes adjacent to the anticodon position. Several other yeast tRNA genes including those coding for $tRNA^{Trp}$ [4], $tRNA^{Ser}_{UCG}$ [5] and $tRNA^{Leu}_{3}$ [6] (Valenzuela, pers. comm.) also contain intervening sequences. In contrast, other yeast tRNA genes such as those coding for $tRNA^{Arg}_{3}$ and $tRNA^{Asp}$ (Beckman, Sakano and Abelson,pers. comm.) and $tRNA^{Ser}_{2}$ and $tRNA^{Ser}_{UCA}$ (Page, Olson and Hall, pers. comm.) lack intervening sequences. The intervening sequences of eukaryotic tRNAs have been shown to be transcribed[7,8] and enzymatically removed to yield mature length tRNAs[7-12].

To determine if the presence of an intervening sequence is a general feature of eukaryotic $tRNA^{Phe}$ genes and to search for structural clues to mechanisms of eukaryotic gene expression, we have isolated and sequenced a Neuro-

spora crassa $tRNA^{Phe}$ gene. We show that this gene contains an intervening sequence similar to those found in yeast $tRNA^{Phe}$ genes. The extent and pattern of sequence conservation in the $tRNA^{Phe}$ intervening sequence, coding region, and surrounding sequences, suggest functional constraints on variation in these genetic regions.

MATERIALS AND METHODS

Total N. crassa tRNA was extracted from germinating conidia[13] as described by Rubin[14]. Purified N. crassa $tRNA^{Phe}$ and $tRNA^{Met}$ were generous gifts of W. E. Barnett of the Oak Ridge National Laboratory. The RNAs were treated with calf intestinal alkaline phosphatase (Boehringer Mannheim), fractionated by electrophoresis through a 10% polyacrylamide gel containing 7M urea[15] and the tRNA fractions were eluted as described for DNA fragments. The RNA was then radioactively labeled using T4 polynucleotide kinase (PL Biochemicals) and [γ-^{32}P-]-ATP (ICN Pharmaceuticals)[16].

Colony hybridizations were performed according to the method of Gruenstein and Hogness[17]. Hybridizations involving DNA transferred from agarose gels to nitrocellulose were performed according to the method of Southern[18] and hybridizations involving DNA transferred from polyacrylamide-agarose gels to diazobenzyloxymethyl paper were performed according to the method of Reiser et al.[19].

Isolation of plasmid DNA[20] and DNA restriction fragments from agarose[20] or polyacrylamide gels[16] was carried out as previously described. To sequence the $tRNA^{Phe}$ gene (coding strand), DNA of plasmid pMF3 (see Results) was digested with Hinf I endonuclease, treated with calf intestinal alkaline phosphatase and fractionated on a 5% polyacrylamide gel. DNA from a triplet band which included the DNA fragment that hybridized to $tRNA^{Phe}$ was isolated and 5' end-labeled for sequencing[21]. To sequence the opposite strand, Hinf I digested pMF3 was 3' end-labeled by filling in with [α-^{32}P]-dGTP, [α-^{32}P]-dCTP, [α-^{32}P]-dATP and [α-^{32}P]-TTP (25 μCi each; 400 Ci/mmole; Amersham) as previously described[22]. After labeling, the DNA was electrophoresed through a 5% polyacrylamide gel and DNA from the triplet band including the DNA fragment which hybridized to $tRNA^{Phe}$ was isolated. The 5' or 3'-labeled Hinf I DNA fragments were subjected to secondary restrictions by Eco RI or Taq I endonucleases and the resulting DNA fragments were resolved by electrophoresis through a non-denaturing 5% polyacrylamide gel or an 8% polyacrylamide gel containing 7M urea (to separate labeled molecules differing in length by only a few nucleotides). The DNA fragments containing the $tRNA^{Phe}$ gene were

identified by size and sequenced according to the methods of Maxam and Gilbert[21].

RESULTS

The N. crassa $tRNA^{Phe}$ gene studied was detected in the collection of pBR322-N. crassa DNA hybrids constructed by Free et al.[23]. Approximately 150 clones containing DNA which hybridized to N. crassa RNA were screened by colony hybridization using radioactively labeled total tRNA, $tRNA^{Phe}$, or $tRNA^{Met}$ of N. crassa as probes. About one-third of the clones contained DNA which hybridized to total tRNA and, of these, the DNA of one clone hybridized specifically to $tRNA^{Phe}$ and that of another hybridized specifically to $tRNA^{Met}$. The $tRNA^{Phe}$ clone contained a 15 kb plasmid. This plasmid was designated pMF3. It consisted of several Pst I fragments of Neurospora DNA inserted at the Pst I site of pBR322[24]. To identify restriction fragments having DNA complementary to $tRNA^{Phe}$, plasmid pMF3 DNA was digested using various restriction endonucleases, and the resulting DNA fragments were fractionated electrophoretically, transferred to nitrocellulose filters or diazobenzyloxymethyl paper and hybridized with ^{32}P-$tRNA^{Phe}$. $tRNA^{Phe}$ hybridized to single pMF3 restriction fragments generated by Pst I. Hind III, Eco RI, Sal I, Bam HI, Xho I, Kpn I, Pvu II, Bgl II, Hinc II, Hpa II, Hae III, Mbo I, Mbo II, Hinf I Taq I or Sau 3A suggesting that one and only one $tRNA^{Phe}$ gene is on this plasmid. Figure 2 illustrates the results of one such experiment which also yielded restriction site data useful in DNA sequence analyses. This

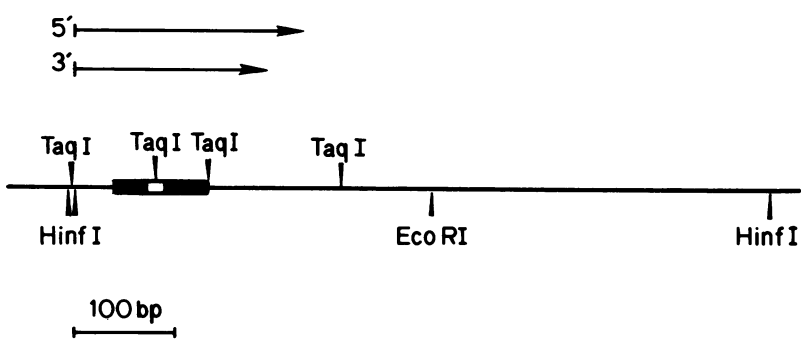

Figure 1. Selected restriction sites in the vicinity of the Neurospora crassa $tRNA^{Phe}$ gene on pMF3. The heavy line marks the position of the tRNA coding region, and the arrows indicate the extent of DNA sequencing relevant to this study. The Taq I site internal to the $tRNA^{Phe}$ gene was identified by DNA sequencing only.

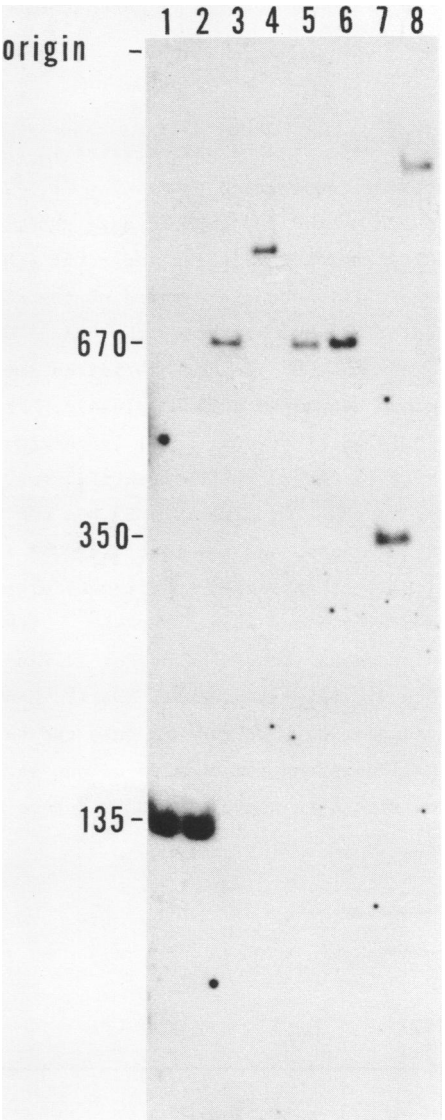

Figure 2. Autoradiogram showing hybridization of ^{32}P labeled N. crassa $tRNA^{Phe}$ to restriction fragments of pMF3. The DNA was digested with Taq I (lane 1), Taq I plus Hinf I (lane 2), Mbo II plus Eco RI (lane 3), Mbo II (lane 4), Mbo II plus Hinf I (lane 5), Hinf I (lane 6), Hinf I plus Eco RI (lane 7) or Bgl II (lane 8). Restriction fragments were electrophoretically separated on a 5% polyacrylamide-0.8% agarose gel and transferred to diazobenzyloxymethyl paper for hybridization. The length (in base pairs) of several DNA fragments which hybridized to $tRNA^{Phe}$ is indicated adjacent to the autoradiogram.

autoradiogram shows that most, if not all, of the $tRNA^{Phe}$ gene is contained in a Hinf I generated DNA fragment approximately 670 bp long (lane 6) which has an Eco RI site about 350 bp from one end (lane 7). The $tRNA^{Phe}$ gene also appears to be within a 135 bp Taq I fragment (lane 1) which is not cut by Hinf I (lane 2).

To determine whether any additional tRNA genes are present on the 5 kb Pst I DNA fragment containing the $tRNA^{Phe}$ gene, this DNA fragment was isolated, digested with Hinf I, Hpa II or Sau 3A and the resulting DNA fragments separated by electrophoresis. The fragments were transferred to nitrocellulose and hybridized with labeled N. crassa total tRNA or $tRNA^{Phe}$. In each case, only the DNA fragment which hybridized to $tRNA^{Phe}$ hybridized to total tRNA. This finding suggests that the $tRNA^{Phe}$ gene is the only tRNA gene on this 5 kb fragment of Neurospora DNA.

Our DNA sequencing strategy was based on the hybridization results obtained with restriction enzyme digests of pMF3 DNA. The Hinf I, Hinf I-Eco RI and Taq I DNA fragments identified from the autoradiogram shown in Figure 2 were isolated (along with several comigrating DNA fragments) and mapped further (data not shown). Finally, the tRNA coding region was sequenced on both strands of the DNA (Figure 1) as described in Materials and Methods. The Neurospora $tRNA^{Phe}$ gene was recognized by its similarity to the yeast $tRNA^{Phe}$ gene[1] and its identity was confirmed by comparison with the sequence of N. crassa $tRNA^{Phe}$ described in the accompanying paper by DeWeerd et al.[25]. The DNA sequence is presented in Figure 3 aligned with the $tRNA^{Phe}$ sequence.

DISCUSSION

We have identified a cloned fragment of Neurospora crassa DNA that contains a $tRNA^{Phe}$ gene. Hybridization analysis suggests that no other tRNA gene is present on the 5 kb cloned fragment. The region of the plasmid containing the $tRNA^{Phe}$ gene was identified by hybridization and was sequenced. The DNA sequence is presented in Figure 3 along with the sequence of $tRNA^{Phe}$ [25] and the nucleotide differences in a yeast $tRNA^{Phe}$ gene. The most striking feature of the gene is a 16 nucleotide sequence starting one base pair after the anticodon position which is not present in mature $tRNA^{Phe}$ [25]. This intervening sequence is located at exactly the same position as those in the yeast tRNA genes that have been studied[1,2,5,6]. However, the intervening sequences of Neurospora and yeast $tRNA^{Phe}$ differ appreciably. The Neurospora $tRNA^{Phe}$ intervening sequence is 2 or 3 nucleotides shorter than those of yeast $tRNA^{Phe}$ genes[1] and the sequences have diverged greatly. Only 50% of the nucleotides

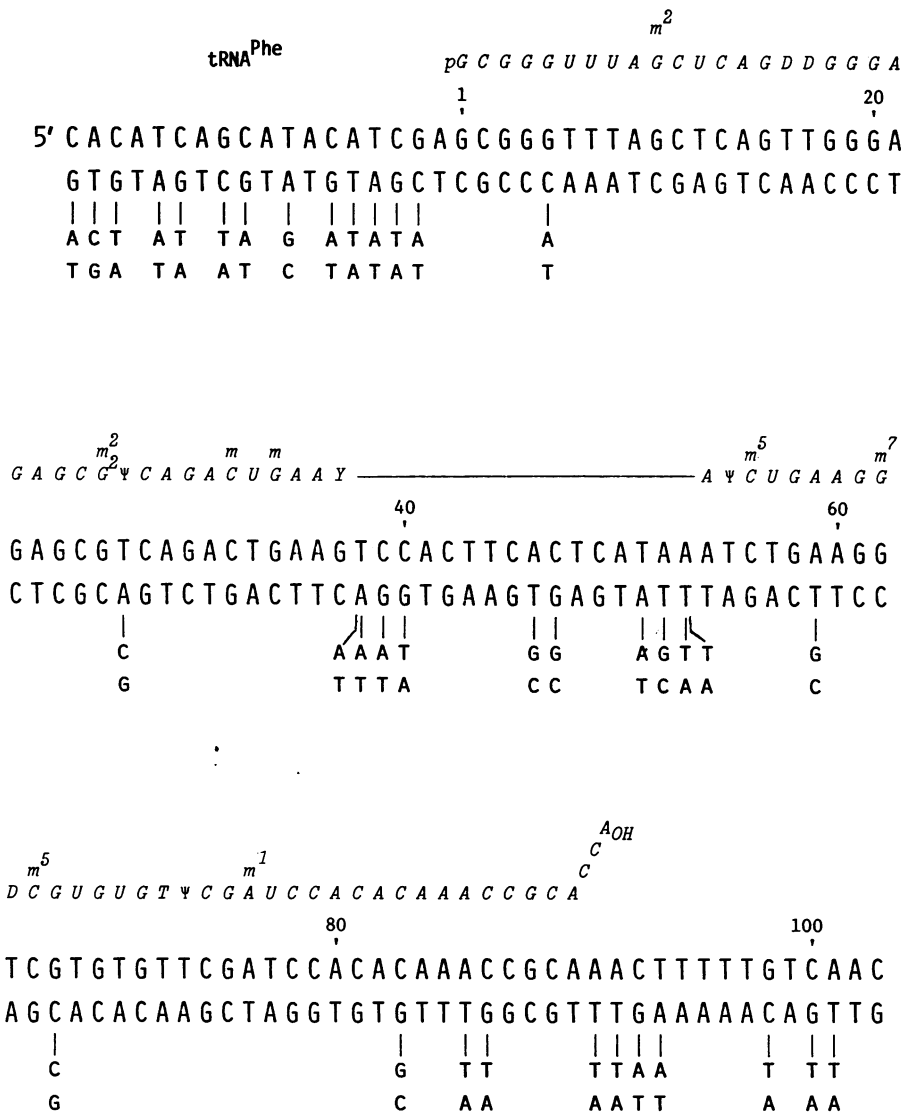

Figure 3. DNA sequence of a Neurospora crassa $tRNA^{Phe}$ gene. The DNA sequence is shown aligned below the N. crassa $tRNA^{Phe}$ sequence[25]. Differences in a yeast $tRNA^{Phe}$ gene sequence (pYP4)[1] are given below the Neurospora sequence.

in the intervening sequences are alike whereas 91% of the base pairs are conserved in the mature $tRNA^{Phe}$ coding regions. This suggests that few structural constraints are imposed by the functions of the intervening sequence. However, although the Neurospora and yeast $tRNA^{Phe}$ intervening sequences have diverged appreciably, they can participate in similar secondary structures. Figure 4 illustrates this point. Both the Neurospora and yeast $tRNA^{Phe}$ precursors can be drawn in secondary structures in which the anticodon region is base-paired with the intervening sequence. The pattern of sequence con-

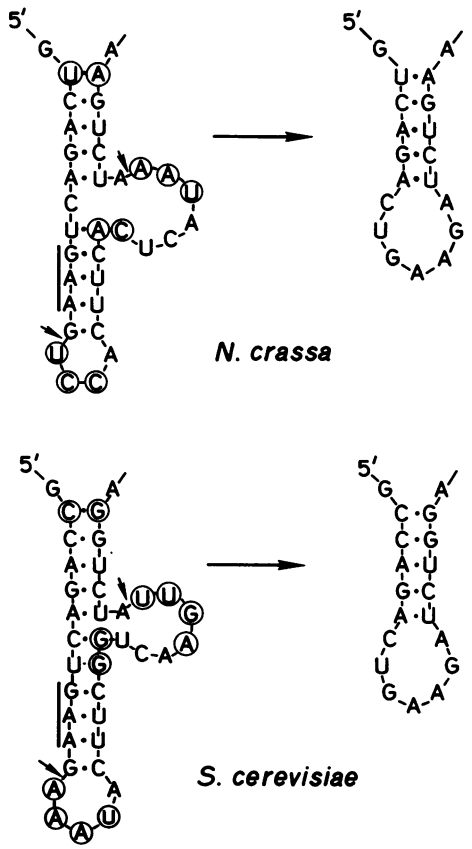

Figure 4. Comparison of proposed secondary structures in $tRNA^{Phe}$ precursors of N. crassa and S. cerevisiae. The anticodon loop regions including the intervening sequences are drawn in possible secondary structures in which the anticodons (position indicated by a vertical line in each precursor) are paired with the intervening sequences. The presumptive cleavage sites are marked with small arrows and the nucleotides which differ in N. crassa and S. cerevisiae are circled.

servation suggests that this anticodon base pairing and the adjacent single stranded loops are important features of $tRNA^{Phe}$ precursors. In contrast,the exact size and nucleotide sequence of the loops are apparently of lesser importance. This conclusion is supported by the similarity among the yeast $tRNA^{Phe}$, $tRNA^{Tyr}$, $tRNA^{Trp}$, $tRNA^{Leu}_{3}$, and $tRNA^{Ser}_{UCG}$ precursors[6]. Although the intervening sequence may have several functions, it seems likely that blockage of anticodon recognition is an essential feature.

The N. crassa $tRNA^{Phe}$ gene shows several interesting features in addition to the intervening sequence. The region following the gene is similar to presumed transcription termination regions following yeast tRNA coding regions[1,2] (Valenzuela, personal communication) and the 5S rRNA coding regions of Xenopus[3], yeast[27,28] and Neurospora (unpublished). All are AT-rich and have a series of A residues in the coding strand, as do several prokaryotic transcription termination regions[29]. The DNA sequence lacks the CCA that is present at the 3' terminus of all mature tRNAs. This is also true of the yeast tRNA genes which have been sequenced but differs from some prokaryotic tRNA genes[30]. The DNA sequence preceding the Neurospora $tRNA^{Phe}$ coding region is very different from the corresponding region of yeast $tRNA^{Phe}$ genes (Figure 3). This is consistent with the variability found preceding yeast tRNA coding regions[1,2] (Valenzuela, personal communication). Transcription studies with cloned $tRNA^{Tyr}$ genes injected into frog oocytes suggest that at least in some cases the DNA segment 5' to the tRNA coding region is transcribed[12]. On the other hand, it has recently been shown that a Xenopus DNA fragment containing a $tRNA^{Met}$ gene and only 22 nucleotides preceding it will support tRNA production in Xenopus oocytes[31]. Whether or not sequences 5' to the transcribed region of tRNA genes play a role in initiation or regulation of transcription is not known.

ACKNOWLEDGEMENTS

This work was supported by grants from the United States Public Health Service (GM-09738) and the National Science Foundation (PCM 77-24333). C.Y. is a Career Investigator of the American Heart Association. E.S. is a predoctoral trainee of the United States Public Health Service. This study was carried out in accordance with the National Institutes of Health Guidelines for Recombinant DNA Research. We thank Stephen Free and Robert Metzenberg for clones from their collection and for advice. We would also like to thank David Setzer for the gift of aminobenzyloxymethal paper and Gerard Zurawski for advice.

REFERENCES

1. Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R. and Rutter, W.J. (1978) Proc. Natl. Acad. Sci. USA 75, 190-194.
2. Goodman, H.M., Olson, M.V. and Hall, B.D. (1977) Proc. Natl. Acad. Sci. USA 74, 5453-5457.
3. Korn, L.J. and Brown, D.D. (1978) Cell 15, 1145-1156.
4. Ogden, R.C., Beckman, J.S., Abelson, J., Kang. H.S. Söll, D. and Schmidt, O. (1979) Cell 17, 339-406.
5. Etchererry, T., Colby, D. and Guthrie, C. (1979) Cell 18, 11-26.
6. Kang, H.S., Ogden, R.C., Knapp, G., Peebles, C.L. and Abelson, J. (1979) ICN-UCLA Symposia. Academic Press, New York, in press.
7. Knapp, G., Beckmann, J.S., Johnson, P.F., Fuhrman, S.A. and Abelson, J. (1978) Cell 14, 221-236.
8. Hopper, A.K., Banks, F. and Evangelidis, V. (1978) Cell 14, 211-219.
9. O'Farrell, P.Z., Cordell, B., Valenzuela, P., Rutter, W.J. and Goodman, H. M. (1978) Nature 274, 438-445.
10. Peebles, C.L., Ogden, R.C., Knapp, G. and Abelson, J. (1979) Cell 18, 27-35.
11. Knapp, G., Ogden, R.C., Peebles, C.L. and Abelson, J. (1979) Cell 18, 37-45.
12. DeRobertis, E.M. and Olsen, M.V. (1979) Nature 278, 137-143.
13. Russell, P.J., Hammett, J.P. and Selker, E.U. (1976) J. Bacteriol. 127, 785-793.
14. Rubin, G.M. (1975) In Methods in Cell Biology. Prescott, D.M., Ed., pp. 45-64, Academic Press, New York.
15. Maniatis, T., Jeffery, A. and van de Sande, H. (1975) Biochemistry 14, 3787-3794.
16. Bennett, G.N., Schweingruber, M.E., Brown, K.D., Squires, C. and Yanofsky, C. (1978) J. Mol. Biol. 121, 113-137.
17. Gruenstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.
18. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
19. Reiser, J., Renart, J. and Stark, G.R. (1978) Biochem. Biophys. Res. Commun. 85, 1104-1112.
20. Selker, E., Brown, K. and Yanofsky, C. (1977) J. Bacteriol. 129, 388-394.
21. Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
22. Selker, E. and Yanofsky, C. (1979) J. Mol. Biol. 130, 135-143.
23. Free, S.J., Rice, P.W. and Metzenberg, R.L. (1979) J. Bacteriol. 137, 1219-1226.
24. Bolivar, F., Rodriquez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) In DNA Insertion Elements, Plasmids, and Episomes. Bukhari, A.I., Shapiro, J.A. and Adhya, S.L. Eds., pp. 686-687, Cold Spring Harbor Press, New York.
25. DeWeerd, B.A., Hecker, L.I., Barnett, W.E. and RajBhandary, U.L. (1980) Nucleic Acids Res. 8, 1023-1032
26. Beckmann, J.S., Johnson, P.F. and Abelson, J. (1977) Science 196, 206-208.
27. Valenzuela, P., Bell, G.I., Masiarz, F.R., DeGennaro, L.J. and Rutter, W.J. (1977) Nature 267, 641-643.
28. Maxam, A., Tizard, R., Skryabin, K.B. and Gilbert, W. (1977) Nature 267, 643-645.
29. Rosenberg, M. and Court, D. (1979) Ann. Rev. Genetics, in press.
30. Smith, J.D. (1976) Progr. Nucleic Acid Res. Mol. Biol. 16, 25-73.

31. Telford, J.L., Kressman, A., Koski, R.A., Grosschedl, R., Muller, F., Clarkson, S.G. and Birnstiel, M.L. (1979) Proc. Natl. Acad. Sci. USA 76, 2590-2594.