Volume 2 number 6 June 1975 Nucleic Acids Research

The 3' terminal oligonucleotide of E.coli 16S ribosomal RNA: the sequence in both wild-type and RNase III⁻ cells is complementary to the polypurine tracts common to mRNA initiator regions

Karen U.Sprague and Joan Argetsinger Steitz

Department of Molecular Biophysics and Biochemistry, Yale University, 333 Cedar Street, New Haven, Connecticut 06510, USA

Received 12 March 1975

ABSTRACT

Application of Sanger techniques to the analysis of the 3' terminal oligonucleotide from E. coli ^{32}P-labelled 16S rRNA yields the sequence AUCACCUCCUUA$_{OH}$. This sequence is identical in RNA isolated from two wild-type strains (MRE600 and E. coli B, SY106) and from a mutant strain (AB301/105) defective in RNase III. Data presented here explains the previous derivation of an incorrect sequence (AUCCUCACUUCA$_{OH}$) by others. The functional significance of complementarity between the 3' terminus of 16S rRNA and poly-purine tracts commonly found in mRNA initiator regions is discussed.

INTRODUCTION

Recently, Shine and Dalgarno[1] suggested that a sequence of 12 nucleotides appearing at the 3' terminus of E. coli 16S ribosomal RNA may participate directly, by base pairing with messenger RNA, in the initiation and termination of protein biosynthesis. Specifically, these authors pointed out that the 3' terminal UUA$_{OH}$ of 16S rRNA is complementary to the termination codons UAA, UAG, and UGA, and that the pyrimidine-rich portion of the terminal sequence could pair with purine-rich regions found 5' to many initiator codons in natural mRNAs.

The 3' terminal sequence deduced by Shine and Dalgarno - AUCACCUCCUUA$_{OH}$ - differs at several critical positions, however, from that proposed by others. The Shine and Dalgarno sequence is based on the application of a sequential degradation technique to unlabelled RNA and in part, sequence data derived by others.[3,7] An alternative isomeric sequence - AUCCUCACUUCA$_{OH}$ - has been obtained using standard Sanger techniques and ^{32}P-labelled RNA.[2,3] Because accurate information about the primary sequence of the 3' terminal oligonucleotide of 16S RNA is essential to current studies of its biological function (Steitz and Jakes, in preparation) and of its secondary conformation in the ribosome (Yuan, Steitz, and Crothers, in preparation), we re-investigated its nucleotide sequence using ^{32}P-labelled RNA and a greater

range of Sanger techniques than were previously applied. Our results agree with those of Shine and Dalgarno[1] and those reported by Ehresmann, Stiegler, and Ebel[38] and by Noller and Herr[4] while this work was in progress.

To examine the requirement for this particular sequence at the 3' end of 16S rRNA in functioning ribosomes, we have compared the 3' terminal oligonucleotide of 16S RNA isolated from wild type E. coli strains with that from a viable mutant strain (AB301/105), defective in one of the ribosomal RNA processing enzymes, RNase III.[5,6,7] We find an identical 3' terminal sequence in 16S rRNA from both strains, compatible with the idea that this region of the rRNA is necessary for direct interaction with mRNA during the initiation of protein synthesis.

MATERIALS AND METHODS

Labelling and Isolation of 16S RNA

Wild type E. coli cells (either MRE600 or a B strain,SY106) were grown in 20 ml of PGM (per liter: 10g peptone, 5g NaCl, 1g glucose, 100 ml 1 M Tris-HCl, pH 7.5) and labelled with 1-2 mCi $^{32}PO_4$ for 2.5 generations in early- to mid-log phase. RNA was extracted by boiling the cells in 1% SDS[8] and was purified by phenol extraction and ethanol precipitation. The individual species were fractionated by sedimentation through a sucrose gradient containing 0.5% SDS, as described by Suzuki and Brown.[9] This procedure routinely yielded 100-200 µg of 16S RNA with a specific activity of 1-4 x 10^6 cpm/µg.

AB301/105 cells (a kind gift of F.W. Studier) were grown in 20 ml of B2ab medium (see Ginsburg and Steitz[10]) in the presence of 1-2 mCi $^{32}PO_4$ for 5 generations. Because these cells appear to contain significant amounts of aberrantly processed rRNA (D. Ginsburg, personal communication) which may not be incorporated into ribosomes, AB301/105 RNA used for sequence analysis was obtained by first isolating ribosomal particles from cells ground with alumina as described by Brownlee and Sanger.[11] RNA was extracted from the ribosomal pellet with boiling 1% SDS, and was fractionated on a sucrose gradient as described above for RNA prepared from whole cells. The gradient profiles indicated that normal proportions of 16S and 23S RNA were present in the ribosomal particles.

Purification of the 16S 3' terminal oligonucleotide

The 16S rRNA was recovered from gradient fractions by precipitation with 2 volumes of ethanol, and was digested (in the presence of sufficient carrier RNA to give a total of 100 µg of RNA) with 10 µl of 1 mg/

ml ribonuclease T_1 in 0.01 M Tris-HCl, pH 7.0, 0.001 M EDTA for 30 minutes at 37°C. The 3' terminal oligonucleotide was isolated by selective binding to dihydroxyboryl-substituted cellulose[12] as described by Rosenberg.[13] The purity of the eluted 3' terminal fragment was checked by electrophoresis on Cellogel at pH 3.5, followed by homochromatography[14] with homomixture C on thin layers of DEAE cellulose (Analtech) or polyethyleneimine (Brinkmann, Cel 300 PEI). The yield of fragment was between 90 and 100% in all preparations. In some experiments, the 3' terminal oligonucleotide produced by digestion with ribonuclease U_2 (10 units/ml for 2 hrs at 37°C) was isolated on dihydroxyboryl cellulose and then purified in a similar fashion.

Alternatively, the 3' terminal oligonucleotide was isolated by the two dimensional electrophoretic method of Dahlberg[15], which takes advantage of the fact that the only T_1 oligonucleotide whose mobility on DEAE paper in 7% formic acid is unaffected by alkaline phosphatase treatment is that originating from the 3' end of the intact molecule. The yield of fragment prepared in this way was approximately 20%.

Sequence analysis

Generally, the methods developed by Sanger and co-workers[14,16] were used in this work. Partial digestion with the 5' exonuclease, spleen phosphodiesterase, was carried out at 37° in 10 µl of the buffer described by Barrell[14] at an enzyme concentration of 0.5 mg (10 units) per ml. Highly purified[17] bovine spleen phosphodiesterase (the generous gift of E.G. Niles) was used for this purpose because, at enzyme concentrations required to cleave 5' C residues, the level of contaminating endonuclease activity in commercial enzyme preparations is unacceptable. Aliquots were removed from the digestion mixture at intervals of 1, 2, and 3 hours; the reaction was stopped by freezing at -60°. Oligonucleotides in the pooled aliquots were fractionated in a two dimensional system, as suggested by Galibert et al.[18]

RESULTS

Isolation of the 3' terminal oligonucleotide produced by ribonuclease T_1 digestion of 16S rRNA

In most experiments, the 3' terminal oligonucleotide was isolated from an RNase T_1 digest of 16S RNA by taking advantage of the affinity of its 2'-3' vicinal hydroxyl groups for dihydroxyboryl-substituted cellulose.[12] The behavior of the eluted 3' terminal fragment in a two dimensional fractionation system is shown in Figure 1. After elution

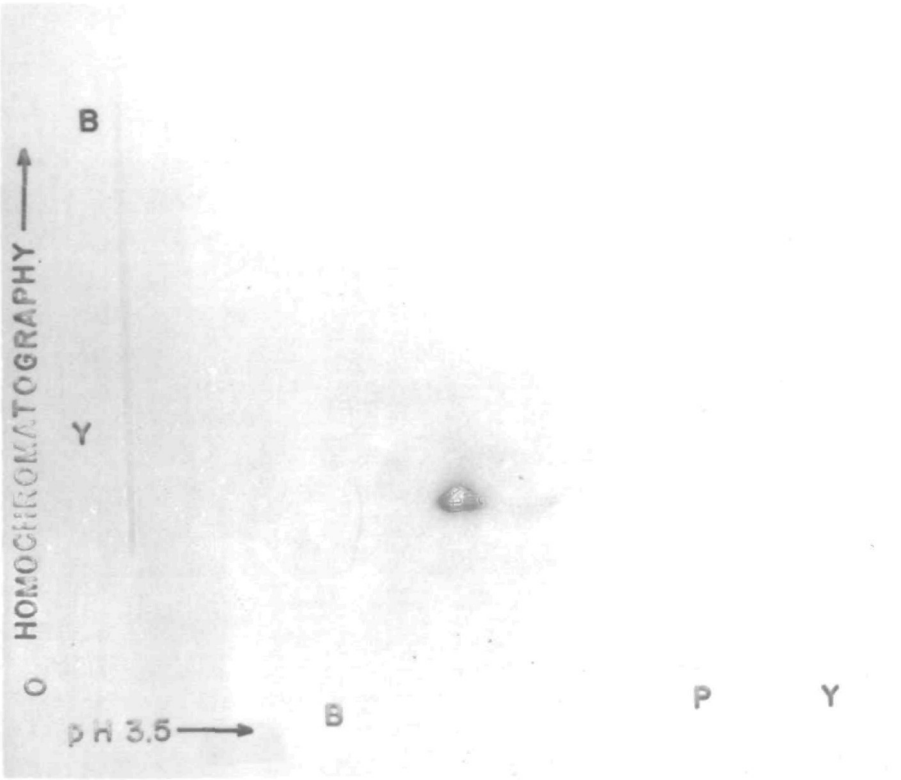


Figure 1. Purification of the 3' terminal T_1 oligonucleotide isolated from E. coli (strain MRE600). The oligonucleotide was recovered from dihydroxyboryl cellulose and was purified in the two dimensional system described in the text. B, P, and Y indicate the positions of the blue, pink and yellow marker dyes.

from the thin layer plate, the purified oligonucleotide was subjected to the following analyses.

Digestion of the oligonucleotide with pancreatic ribonuclease

The identity and yields of the products of pancreatic ribonuclease digestion shown in Table 1 are in agreement with data reported by others.[2,19] Taken together with the chromatographic mobility of the fragment (Figure 1), which suggests a length of 10-12 nucleotides, they establish the composition: AU, AC, C_4, U_3, X_{OH} for the 3' terminal RNase T_1 oligonucleotide.

Digestion with pancreatic ribonuclease after carbodiimide blocking

Both proposed sequences for the 16S 3' terminal oligonucleotide should yield 4 identical products (AC, C, AUC, UC) after pancreatic

Table 1

Products obtained by pancreatic ribonuclease digestion of the 16S rRNA 3' terminal T_1 oligonucleotide

Oligonucleotide	Relative moles Pi/oligonucleotide
AU	2.2
AC	2.0
C	4.0
U	3.0

ribonuclease digestion of the carbodiimide-blocked fragment. A fifth product, diagnostic for the correct sequence, would be UUA_{OH} in the Shine and Dalgarno[1] sequence, or CUU in the Santer and Santer[3] sequence. The results we obtained are shown in Figure 2. The products were

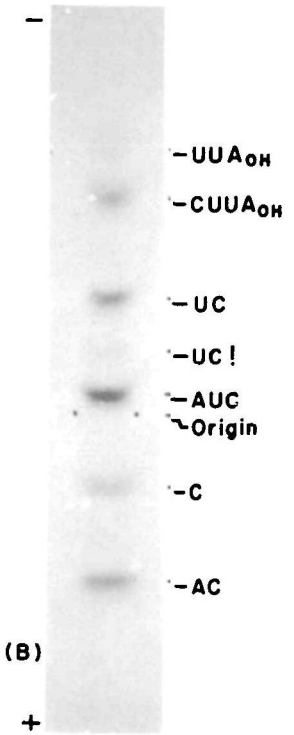


Figure 2. Products obtained by pancreatic ribonuclease digestion of the 3' terminal T_1 oligonucleotide blocked with carbodiimide. Fractionation was performed according to Barrell.[14] (B) marks the position of the blue marker dye.

identified as AC, C, AUC, UC, UUA_{OH}, and $CUUA_{OH}$ by digestion of each with pancreatic ribonuclease and with snake venom phosphodiesterase after removal of the carbodiimide group.[14] The identity of the UUA_{OH} oligonucleotide was verified by comparison of its electrophoretic mobility with that of a carbodiimide-derivatized UUA_{OH} marker prepared from the 3' terminus of B. mori 18S rRNA (Sprague and Kramer, manuscript in preparation).

Santer and Santer[2] reported that pancreatic ribonuclease digestion of the carbodiimide-blocked T_1 oligonucleotide produces UUC, rather than UUA_{OH}. In repeated trials, we found variable yields of the partial digestion product, $CUUA_{OH}$. This observation suggests that confusion may have arisen previously because of incomplete cleavage between the C residue and the two adjacent carbodiimide-modified U residues in this part of the molecule. The resulting $CUUA_{OH}$ oligonucleotide can be distinguished from UUC only by subsequent digestion with snake venom phosphodiesterase, as performed here. Examples of such incomplete cleavage of carbodiimide-blocked oligonucleotides have also been

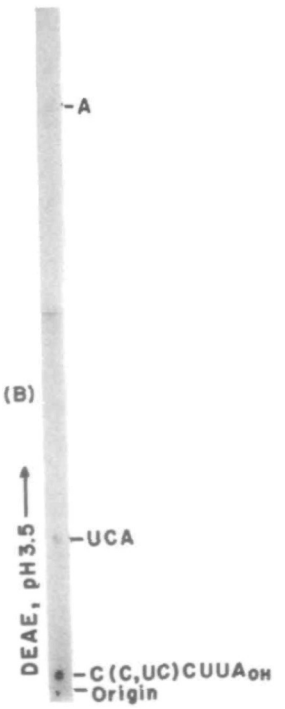


Figure 3. Products obtained by digestion of the 3' terminal T_1 oligonucleotide with ribonuclease U_2. (B) marks the position of the blue marker dye.

observed by others in the analysis of E. coli 5S RNA[20] and the T_7 bacteriophage early messenger RNAs (R. Kramer, personal communication).

Digestion with ribonuclease U_2

The two proposed sequences predict that treatment of the 16S terminal T_1 oligonucleotide with ribonuclease U_2 should yield a distinct set of products in each case: A, UCA, and $CCUCCUUA_{OH}$ from the Shine and Dalgarno[1] sequence; A, UCCUCA, and $CUUCA_{OH}$ from the Santer and Santer[3] sequence. Figure 3 shows an autoradiograph of the U_2 products we obtained. These were subsequently identified by hydrolysis with alkali and with snake venom phosphodiesterase, or by pancreatic ribonuclease digestion after blocking U residues with carbodiimide. The three oligonucleotides are those expected from the sequence proposed by Shine and Dalgarno.[1]

This analysis, in combination with those described in the preceding sections, establishes the following sequence for the 3' terminus of 16S rRNA: $AUCAC(C,UC)CUUA_{OH}$.

Partial digestion with spleen phosphodiesterase

In order to resolve the (C,UC) ambiguity in the above sequence, the largest product of ribonuclease U_2 cleavage, $C(C,UC)CUUA_{OH}$, was isolated and subjected to partial digestion with the 5' exonuclease, spleen phosphodiesterase. The resulting oligonucleotides were identified in two ways: 1) by the characteristic alteration in electrophoretic and chromatographic mobility accompanying removal of a single 5' nucleotide, and 2) by subsequent analysis using alkaline hydrolysis and complete digestion with snake venom phosphodiesterase. The results shown in Figure 4 demonstrate that the sequence of the 3' terminal U_2 oligonucleotide is $CCUCCUUA_{OH}$.

Comparison of the 16S rRNA 3' terminal oligonucleotide from wild-type and AB301/105 cells

The 3' terminal fragment produced by ribonuclease T_1 cleavage of AB301/105 16S rRNA was indistinguishable from the corresponding fragment derived from E. coli B or MRE600 cells. The electrophoretic and chromatographic properties of the terminal oligonucleotides did not differ, and digestion with pancreatic ribonuclease after carbodiimide blocking yielded identical products.

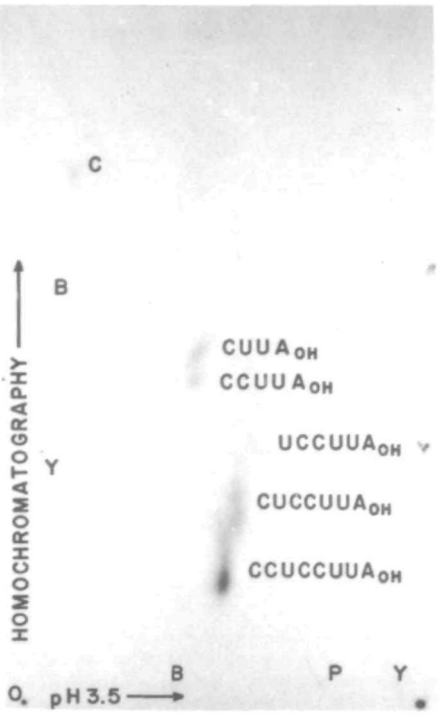


Figure 4. Partial cleavage products obtained by digestion of the 3' terminal U_2 oligonucleotide with spleen phosphodiesterase. B, P, and Y mark the positions of the blue, pink, and yellow marker dyes. Further analysis of these oligonucleotides with alkali and snake venom phosphodiesterase yielded the molar ratios shown below.

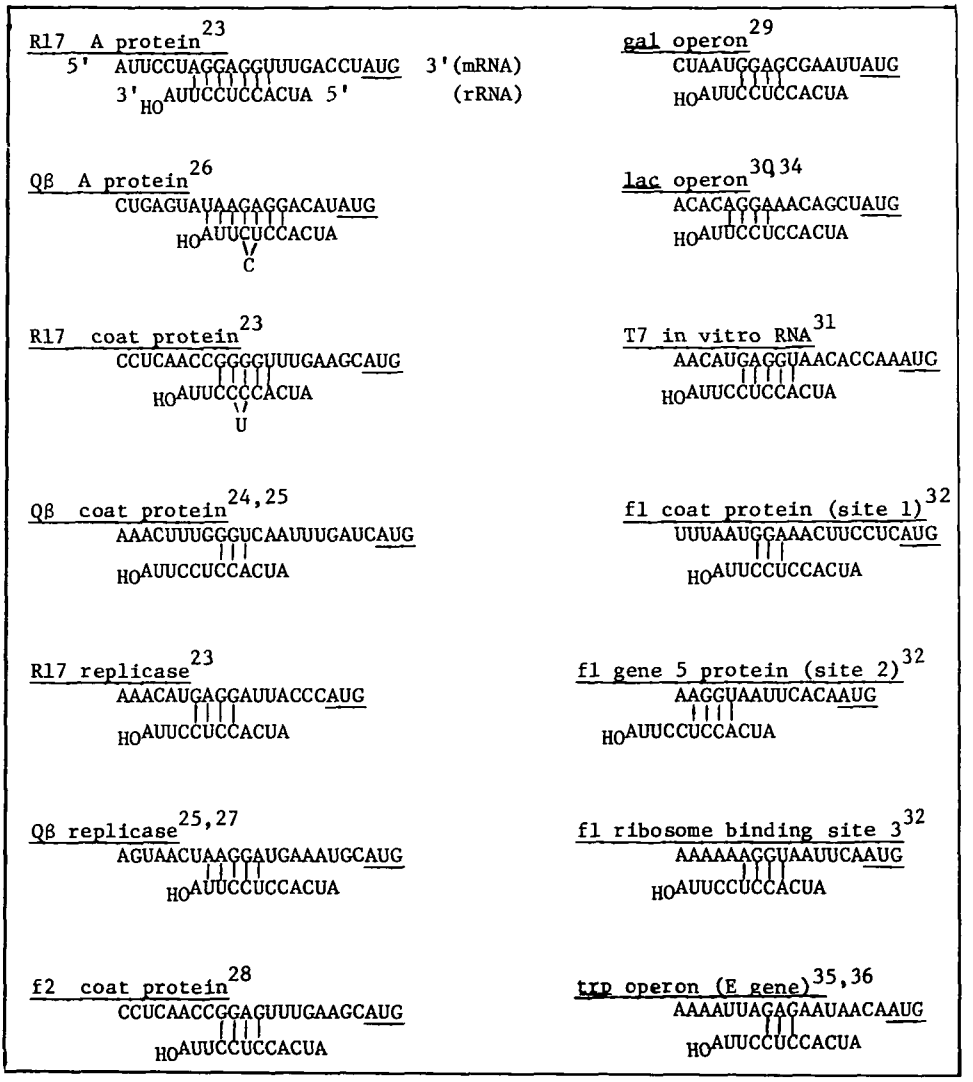
oligonucleotide	alkali Up/Cp	snake venom phosphodiesterase pU/pC	pU/pA
CCUCCUUA$_{OH}$	0.80	1.0	3.1
CUCCUUA$_{OH}$	1.0	1.2	2.9
UCCUUA$_{OH}$	1.5	1.1	1.8
CCUUA$_{OH}$	1.0	2.1	2.0
CUUA$_{OH}$	1.8	no C	1.7

DISCUSSION

In agreement with the results of Shine and Dalgarno[1] and of Noller and Herr[4], we find the sequence AUCACCUCCUUA$_{OH}$ at the 3' terminus of 16S rRNA from E. coli. Discrepancies between this sequence and that proposed by Santer and Santer[2,3] are unlikely to be due to strain differences, since we have examined 16S rRNA from both E. coli strains

Table 2

Complementarity Between Ribosome Binding Sites and The 3' Terminus of E. coli 16S rRNA



The base-paired structures shown are the most stable, calculated according to Tinoco et al.[22] and assuming stacking of the polypurine regions.

used previously and have found that the 3' terminal T_1 oligonucleotides are identical. We believe that errors in the Santer and Santer sequence may be attributed to the difficulties inherent in applying Sanger techniques to the sequence analysis of pyrimidine-rich oligonucleotides. In particular, problems arose because of: 1) mis-identification of the partial cleavage product $CUUA_{OH}$ arising from pancreatic ribonuclease digestion of carbodiimide-blocked RNA, 2) the absence of data obtained from digestion with ribonuclease U_2, and 3) the difficulty of interpreting quantitative data derived from partial digestion with snake venom phosphodiesterase of pyrimidine-rich oligonucleotides.

The possibility that portions of the 16S rRNA may play critical roles in the initiation and termination of protein synthesis was first suggested by Shine and Dalgarno[1]. Purine-rich regions appearing about 10 nucleotides 5' to the initiator codon are indeed among the few identifiable common features of prokaryotic ribosome binding sites (Table 2). Consistent with the functional relevance of these purine-rich stretches is the observation that ribosome binding to natural messenger RNA is competitively inhibited by certain random co-polymers of A, U, and G which are rich in A and G; other synthetic polynucleotides and the trinucleotide AUG are much less efficient competitors.[21] Moreover, T_1 RNase fingerprints of ribosome-protected initiator regions from a variety of prokaryotic messenger RNAs invariably contain higher than expected yields of G and AG (J.A. Steitz, unpublished observations), indicative of a high affinity of ribosomes for purine-rich nucleotide sequences. Current evidence suggesting that the 3' end of 16S rRNA may lie close to the mRNA binding site in the functional ribosome is cited by Shine and Dalgarno.[33]

Finally, our finding that ribosomes isolated from cells defective in the processing of ribosomal RNA nonetheless contain 16S RNA with a normal 3' terminal sequence argues for the importance of this region of the RNA in some aspect of ribosome assembly or function.

Clearly, the data presented above supports, but does not critically test, the hypothesis proposed by Shine and Dalgarno.[1] For this purpose, functional assay of ribosomes containing modified 16S RNA, or direct isolation of an mRNA-16S RNA complex (Steitz and Jakes, in preparation) will be required.

ACKNOWLEDGMENTS

We would like to thank A.W. Senear and R.A. Kramer for their help with the preparation and sequence analysis of radioactive 16S rRNA, and we are particularly grateful to M. Santer for valuable discussions throughout the course of this work. Financial support for this project came from a National Institutes of Health Grant (No. AI-10243).

REFERENCES

1. Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. U.S. 71, 1342-1346.

2. Santer, M. and Santer, U.V. (1972) Biochemistry 11, 786-791.

3. Santer, U.V. and Santer M. (1972) FEBS Letters 21, 311-314.

4. Noller, H.F. and Herr, W. (1974) Mol. Biol. Rep. 1, 437-439.

5. Kindler, P., Keil, T.U. and Hofschneider, P.H. (1973) Molec. Gen. Genetics 126, 53-59.

6. Dunn, J.J. and Studier, F.W. (1973) Proc. Natl. Acad. Sci. U.S. 70, 1559-1563.

7. Nikolaev, N., Silengo, L. and Schlessinger, D. (1973) Proc. Natl. Acad. Sci. 70, 3361-3365.

8. Kramer, R.A., Rosenberg, M. and Steitz, J.A. (1974) J. Mol. Biol. 89, 767-776.

9. Suzuki, Y. and Brown, D.D. (1972) J. Mol. Biol. 63, 409-429.

10. Ginsburg, D. and Steitz, J.A. (1975) J. Biol. Chem., in press.

11. Brownlee, G.G. and Sanger, F. (1967) J. Mol Biol. 23, 337-353.

12. Rosenberg, M. and Gilham, P.T. (1971) Biochem. Biophys. Acta 246, 337-340.

13. Rosenberg, M. (1974) Nucleic Acids Research 1, 653-671.

14. Barrell, B.G. (1971) Procedures Nucleic Acid Res. 2, 751-779.

15. Dahlberg, J.E. (1968) Nature 220, 548-552.

16. Sanger, F., Brownlee, G.G. and Barrell, B.G. (1965) J. Mol. Biol. 13, 373-398.

17. Bernardi, A. and Bernardi, G. (1968) Biochim. Biophys. Acta 155, 360-370.

18. Galibert, F., Sedat, J. and Ziff, E. (1974) J. Mol. Biol. 87, 377-407.

19. Ehresmann, C., Fellner, P. and Ebel, J.-P. (1971) FEBS Letters 13, 325-328.

20. Brownlee, G.G., Sanger, F. and Barrell, B.G. (1968) J. Mol. Biol. 34, 379-412.

21. Revel, M. and Greenshpan, H. (1970) Eur. J. Biochem. 16, 117-122.

22. Tinoco, jun. I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40-41.

23. Steitz, J.A. (1969) Nature 224, 957-964.

24. Hindley, J. and Staples, D.H. (1969) Nature 224, 964-967.

25. Steitz, J.A. (1972) Nature New Biol. 236, 71-75.

26. Staples, D.H., Hindley, J., Billeter, M.A. and Weissmann, C. (1971) Nature New Biol. 234, 202-204.

27. Staples, D.H. and Hindley, J. (1971) Nature New Biol. 234, 211-212.

28. Gupta, S.L., Chen, J., Schaefer, L., Lengyel, P. and Weissman, S.M. (1970) Biochem. Biophys. Res. Comm. 39, 883-888.

29. Musso, R., DeCrombrugghe, B., Pastan, I., Sklar, J., Yot, P. and Weissman, S.M. (1974) Proc. Natl. Acad. Sci. U.S. 71, 4940-4944.

30. Maizels, N.M. (1973) Proc. Natl. Acad. Sci. U.S. 70, 3585-3589.

31. Arrand, J.R. and Hindley, J. (1973) Nature New Biol. 244, 10-12.

32. Pieczenik, G., Model, P. and Robertson, H.D. (1974) J. Mol. Biol. 90, 191-214.

33. Shine, J. and Dalgarno, L. (1975) Nature 254, 34-38.

34. Maizels, N. (1974) Nature 249, 647-649.

35. Bronson, J.J., Squires, C. and Yanofsky, C. (1973) Proc. Natl. Acad. Sci. U.S. 70, 2335-2339.

36. T. Platt, personal communication.

37. Fellner, P., Ehresmann, C., Stiegler, P., and Ebel, J.-P. (1972) Nature New Biol. 239, 1-5.

38. Ehresmann, C., Stiegler, P., and Ebel, J.-P. (1974) FEBS Letters 49, 47-48.