Nucleotide sequence of the *rrnG* ribosomal RNA promoter region of *Escherichia coli*

Wei-Fang Shen, Craig Squires and Catherine L.Squires

Department of Biological Sciences, Columbia University, New York, NY 10027, USA

Received 3 February 1982; Revised and Accepted 19 April 1982

ABSTRACT

The primary structure of the promoter region for a ribosomal RNA transcription unit (*rrnG*) of *Escherichia coli* K12 has been determined. The sequence was obtained from a 1.5 kbp EcoRI fragment derived from the hybrid plasmid pLC23-30. This fragment contains 455 bp preceding P1 of the *rrnG* promoter region and 675 bp of the *rrnG* 16S RNA gene. The sequence before the *rrnG* promoter region contains an open reading frame (ORF-BG) followed by a possible hairpin structure that resembles other known transcription terminators. The sequence of the *rrnG* promoter region is similar but not identical to that of *rrnA* and *rrnB*. Several minor differences between the sequences of the 16S RNA genes of *rrnG* and *rrnB* were also noted. In addition, sequences were found that could generate special structures involving the promoter regions of *rrn* loci. Such structures are described and their possible involvement in the regulation of ribosomal RNA synthesis is discussed.

INTRODUCTION

The *Escherichia coli* K12 genome contains seven distinct transcription units that code for ribosomal RNA (Fig. 1). These seven loci have been mapped in part by heteroduplex analysis of *E. coli* gene bank clones and several hybrid λ -transducing phage (1). The assignment of the rRNA locus carried on the Clarke-Carbon plasmid pLC23-30 as *rrnG* (at 56 min on the *E. coli* chromosome) is based on the observation that an *rrn* locus exists on the λ -*phe415* phage isolated by Zurawsky and Brown (2) and that heteroduplex analyses show that pLC23-30 and λ -*phe415* are homologous on both sides of the *rrn* locus (3). One of the seven loci (*rrnB*) has been sequenced in its entirety (4). In addition, regions of interest in other *rrn* loci have been sequenced, including the promoter regions of *rrnA*, *rrnD*^{*}, *rrnE* and *rrnX* [a hybrid locus in which the promoter region is probably derived from *rrnH* of *E. coli* K12 (3,7)] (5,6). The promoter regions of *rrnA* and *rrnB* are identical. The promoter regions of *rrnD*^{*}, *rrnE* and *rrnX* are also quite similar but differ substantially from those of *rrnA* and *rrnB*.

We have determined the (*rrnG*) promoter region sequence carried on plasmid pLC23-30 using "shotgun" cloning into the single stranded DNA bacteriophage M13mp7 (8). We have noted similarities and differences between the *rrnG* and the *rrnB* promoter regions which may be helpful in establishing common structural features involved in the mechanism of ribosomal RNA promotion.

MATERIALS AND METHODS**Source of DNA sequenced**

The source of the DNA sequenced in this work was the Clarke-Carbon plasmid pLC23-30 which has been shown to carry genes from the 56 minute region of the *E. coli* chromosome. A preliminary restriction map of pLC23-30 is shown in Figure 1.

Strains, phage and plasmids

Strain 71.18 ($\Delta lac-proAB supE thi / F'lec^q \Delta M15 proA^+B^+$) (9) and phage M13mp7 (8) were obtained from J. Messing. Different isolates of the Clarke-Carbon plasmid pLC23-30, in *E. coli* strain JA200 (C600 $\Delta trpE5 recA thr^- leu^- lacY^- / F^+$) (10), were obtained from Louise Clarke and from Masayasu Nomura. Strains containing pLC23-30 were grown in L-broth in the presence of colicin E1 (11) to keep the recombinant colE1 plasmid under selective pressure.

DNA sequencing

Techniques and methods that we used for preparing DNA fragments, "shotgun" cloning into the M13mp7 vector system and sequencing by the Sanger dideoxy NTP chain termination method are described elsewhere (12). Computer programs for sequence analysis were obtained from R.

Figure 1. Genetic and physical maps of the plasmid pLC23-30. A. Genetic map of the *E. coli* K12 chromosome showing the seven *rrn* loci. B. Genetic map of the plasmid pLC23-30. The heavy horizontal line indicates the extent of the ColE1 (vector) DNA sequences. The light horizontal line represents sequences derived from *E. coli* K12. Double lines indicate the locations of the rRNA genes of *rrnG*.

I. Region of pLC23-30 DNA which is not homologous with λ ph415 DNA. This region is either an insertion in pLC23-30 (as suggested in ref. 3) or a deletion in the phage DNA.

C. Restriction sites in the plasmid pLC23-30. These maps are aligned with the genetic map immediately above. Single HpaI and XbaI sites are shown on the top line. The sequence determined in this work was derived from the 1.5 kbp EcoRI fragment that spans the promoter region.

ii. There is an additional PstI restriction site 500 bp from one of the two sites in the *E. coli* segment of the plasmid.

iii. PvuII map is incomplete. Only those PvuII sites around *rrnG* have been mapped.

Staden (13). A computer program called "basepr" (B. Bush and C. Squires, unpublished) was used to analyze nucleotide sequences for possible secondary structures. The "shotgun" cloning of HpaII, Sau3A, TaqI and RsaI+HaeIII fragments derived from the 1.5 kbp EcoRI fragment was accomplished with DNA isolated by agarose gel electrophoresis (Sea Plaque, Marine Colloids). We have used a rapid hybridization method (14) to identify the orientation of some of the sequences cloned into M13. Clones are designated plus (+) if they hybridize with 16S RNA or if they are derived from the same strand that codes for 16S RNA.

RESULTS AND DISCUSSION

Determination of DNA sequence

The DNA sequence of the fragment spanning the beginning of the *rrnG* transcription unit was determined from a number of smaller fragments cloned into the vector M13mp7. We examined M13 clones of HpaII, Sau3A, TaqI and RsaI plus HaeIII fragments derived from the 1.5 kbp EcoRI fragment from pLC23-30 (Fig. 2). These clones are designated by letters representing the restriction endonuclease (H, S, T and RH for HpaII, Sau3A, TaqI and RsaI plus HaeIII respectively), a number representing the position of the fragment (the fragments are numbered from left to right in Fig. 2) and a (+) or (-) designating the strand orientation of the clone. Thus, S4(+) is the Sau3A clone that contains the early *rrnG* 16S sequence (nucleotide numbers 15 to 269 of the mature 16S RNA sequence) and that hybridizes with 16S rRNA.

The left and right ends of the sequence were determined from "wrap-around" sequences obtained by circularizing the 1.5 kbp EcoRI fragment with ligase before "shotgun" cloning HpaII fragments. This procedure gave clones that contained the two end fragments (H1 and H8 in Fig. 2), fused at the EcoRI site. Clones of both orientations [i.e. sequence, H8-H1(+); and antisequence, H1-H8(-)] were obtained. The sequences obtained for H8 are identical to 16S sequences found in *rrnB* (4) identifying that segment as the right end of the 1.5 kbp fragment. A similar "wrap-around" clone containing fragments T1-T6(-) was also found although no attempt had been made

Figure 2. Strategy diagram showing DNA sequences obtained.

to ligate the two EcoRI ends before the TaqI cleavage and subsequent "shotgun" cloning. Presumably the two TaqI-EcoRI fragments (T1 and T6) were joined at the EcoRI site during the cloning process.

Sequences were obtained that span all of the restriction sites shown in Figure 2 except for the Sau3A site between S4 and S5 and the TaqI site between T4 and T5. These fragments were assumed to be contiguous at the two sites in question because adjacent sequences from *rrnG* are identical to those from *rrnB* in both regions. The DNA sequence obtained is shown in Fig. 3. Entire sequences for both the (+) and (-) orientations were obtained except for the (+) strand in the region between #418-656 (which spans the promoter region) and the (-) strand between

```

      10      20      30      40      50      60
TTCATTAACC GTATCGATGA AGTGGTGGTC TTCCATCCGC TGGGTGAACA GCACATTGCC
AGAAATGCTG GGCTCGGCAG AGATGGTTGA GTTGGCTGAA GCGAAGCTAC ATGGCGAAGA

      70      80      90     100     110     120
TGGATTGGCG AGATTCAGTT GAAACGCTCG TACAAACGTC TGGGAAGAAG TGGTTATGAA
TGTTTCTCTG GATGCACTAA AACGTATCCT ACGCCCGTGG TTAAGATGA AAGAGCCGCC

      130     140     150     160     170     180
ATCCACATTT CTGACGAGGC GCTGAAACTG CTGAGCGAGA ACGGTTACGA TCCGGTCTAT
AGATACGGTT GTATTGGGTT GCACCCATTT CCTCTACTA CAAGAAGAAC TGTTACAAGT

      190     200     210     220     230     240
GGTGCACGTC CTCTGAAACG TGCAATTCAG CAGCAGATCG AAAACCCGCT GGCACACGAA
GCTGCCAGAG GGAACCCGCG TGSTGGATTG TGGCGCAGCG ATTGCTCGCC GAACGGCCTG

      250     260     270     280     290     300
ATACTGTCTG GTGAATTGGT TCCGGGTAAA GTGATTCGCC TGGGAAGTAA TGAAGACCGG
GTTGTTAGAA CATGAAGCCC CGGATGCAAA ATCTGCCGAT GCGAATATTG CCTTTTGTAT

      310     320
ATTGTCGCCG TCCAGTAAAT
GGCAATGACG CCAGGAGCTG

ORF-BG ter
      330     340     350     360
GATAAACGGA GCCCTTCGGG GCTCGTTTTT GTCTATAAGT
GCTCGAAGAA GCGCGGCTTA CAGCGTTACG GCTTCGAAAC

      370     380     390     400     410     420
TAGACGGAAA AGACTATATT TAAGATGTTT TGCTTGAAAA GTGAGCGAAC GATAAAGTTT
GCTCGAAAAA CTGGCAGTTT TAGGC76ATT TGGT76AATG TTGCGCGGTC AGAAAATTAT

      430     440     450     460     470     480
TTATATTTTT CGCTTGTCAG GCGGGAATAA CTCCCTATTA TCGCCACCA CTGACACGGA
TTAAATTTTC CTCTTGTCAG GCGGGAATAA CTCCCTATTA TCGCCACCA CTGACACGGA

      490     500     510     520     530     540
ACAACGGCAA ACACGCCGCC GGCTCGCCCG GGTTCCTCTG AGAATCTCAA CAGAGAAAAA
ACAACGGCAA ACACGCCGCC GGCTCGCCCG GGTTCCTCTG AGAATCTCAA CAGAGAAAAA

      550     560     570     580     590     600
CAAA6AAATG CTTGACTCTG TAGCGGGAAG GCGTATTATG CACACCGCCG GCGCGCTGAG
AAAAT7AAAT GCTTGACTCT GTAGCGGGA GCGGTATTAT GCACACCCCG GCGCGCTGAG

      610     620     630     640     650     660
AAAAAGCGAA GCGGCACCTG TCTTTAACAA TTTATCAGAC AATCTGTGTG GGCACCTCGAA
AAAAAGCGAA GCGGCACCTG TCTTTAACAA TTTATCAGAC AATCTGTGTG GGCACCTCGAA

      670     680     690     700     710     720
GATACGGATT CTTAAGCTCG CAAGACGAAA AATGAATACC AAGTCTCAAG AGTGAACACG
GATACGGATT CTTAAGCTCG CAAGACGAAA AATGAATACC AAGTCTCAAG AGTGAACACG

```

Figure 3. Sequences surrounding the promoter regions of *rrnG* (top lines) and *rrnB* (bottom lines). The two DNA sequences are aligned at the right hand EcoRI site. Conserved sequences are underlined. Differences and other features discussed in the text are italicized.

#659-690 and #1039-1239. Examination of a large number of clones in M13mp7 revealed that not all of the subfragments were cloned. Larger fragments were obtained less frequently, as observed by other workers (15). In addition, we found that it was difficult to obtain clones that span either P1 and/or P2 (H5, H6, S3 and T4). However, several clones were obtained that had the minus (-) orientation, allowing completion of the sequence. It is interesting to note that some of the missing clones [H5(+), H6(+), S3(+) and T4(+)] would orient the *rrnG* promoter so that any

resulting transcript would be directed towards, and perhaps through, the origin of M13 replication (16). A strong promoter in this orientation may interfere with M13 replication.

Comparison of the *rrnG* and *rrnB* sequences

The ribosomal transcription units sequenced to date all have two tandem promoters. Each promoter has an A+T rich section followed by the usual -35 and -10 regions (17). Transcript starts are either seven (P2) or eight (P1) nucleotides after the -10 sequence (6,18). The spacing of the two promoters is similar for all of the loci that have been sequenced. A highly conserved region 5' to the 16S structural sequence is found in all loci that have been sequenced and is involved in precursor rRNA structures necessary for processing (19). The 16S structural sequence is also highly conserved (1,3) although some heterogeneity has been noted (20).

Figure 3 shows the sequences of the *rrnG* and *rrnB* (4) promoter regions aligned at the right-hand EcoRI site. The two sequences are dissimilar for the first 432 nucleotides, but each contains the 3' end of an open reading frame in this region. Thereafter, the *rrnG* and *rrnB* sequences are similar although not identical. The differences between the two sequences are discussed in greater detail below.

Open reading frames. The open reading frame found before *rrnG* (ORF-BG) occupies the first 315 nucleotides of the *rrnG* sequence in Fig. 3. Preliminary observations have indicated that an RNA about 2 kbp long is transcribed *in vivo* from the *E. coli* chromosome and ends in the region preceding the *rrnG* promoter (S. Aksoy, data not shown). Additional sequencing in the region to the left of the 1.5 kbp EcoRI fragment has indicated that ORF-BG is extensive (W.-F. Shen, data not shown). We have also noted a terminator-like structure (*324-349) following the translation termination codon (*316-318) of ORF-BG (Fig. 4). This structure is similar to that of other strong rho-independent transcription terminators (17).

These observations, taken together, suggest the presence of a separate transcription unit before the *rrnG* promoter region. Work is currently in progress to demonstrate the precise end of the *in vivo* ORF-BG transcript. An open reading frame (ORF-BB) of at least 867 nucleotides is present in the sequence preceding *rrnB*; however, no terminator structure is evident at its end (4). *In vitro* transcription experiments in which a lambda phage promoter is fused to ORF-BB show that transcription does not terminate but continues through the promoter region of *rrnB* (21). These *in vitro* results, if also true *in vivo*, suggest a difference in the role that the preceding genes play in the expression of *rrnG* and *rrnB*.

Promoter regions and 16S precursor structure. The A+T rich regions preceding the P1 -35 regions of *rrnG* and *rrnB* (*408-432) have quite different sequences but both have the same high A+T content (21/25 base pairs). Eleven nucleotides of the *rrnG* promoter region differ from those of the *rrnB* sequence for the region between the -35 sequence of P1 and the 16S gene (*433-761). None of these differences involves -35 or -10 sequences and the P1 -35 and -10 sequences of both loci are the same distance from the 16S gene. However, *rrnB* has one extra nucleotide between P1 and P2 and in *rrnG* there is an additional G (*587) seven nucleotides after

Figure 4. Proposed secondary structure of ORF-BG transcription terminator. This structure is represented in the RNA form because of its proposed termination function.

the P2 -10 sequence which misaligns the P2 regions of the two loci. All P2 *in vitro* transcripts that have been studied initiate at or near a C seven nucleotides following the P2 -10 sequence (6,18). It will be interesting to see if the *in vitro* transcript from *rrnG* also starts at the seventh nucleotide corresponding to the G at position #587. The A+T rich region (#521-550) preceding the P2 -35 region is also different in *rrnG* but the two regions have about the same A+T content (*rrnG* has 20/30 bp vs. 19/30 for *rrnB*). The largest difference in this region is the TCTCAA (*rrnG*) vs. CTCCGG (*rrnB*) at positions #525-530. This and other differences in the promoter region sequence are discussed below with reference to a possible secondary structure for the *rrnG* and *rrnB* promoter regions.

Another difference between *rrnG* and *rrnB* occurs eleven nucleotides (#751, Fig. 3) before the 5' end of the mature 16S sequence. There is an A at this position in five loci [*rrnG*, *rrnB* from pLC34-34(C. Squires, unpublished), *rrnD*^{*} and *rrnX* (6) and *rrnE* (5)]. Both *rrnA* (5) and *rrnB* from *λdrif*^d18 (4) have a G at this position. This difference results in a A:U to G:U change in the proposed secondary structure of precursor rRNA (19). This change should have a very small effect on the formation or stability of the 16S precursor structure. However, the observation of Kingston and Chamberlin that *rrnB* *in vitro* transcripts terminate at secondary structure in this region (22) suggests regulatory intricacies that might be affected by minor changes in sequence.

16S structural sequences. Six differences were found between the mature 16S sequences of *rrnG* and *rrnB* for the 675 nucleotides shown in Fig. 3. Some of these differences are present in RNA sequences obtained from mature 16S rRNA (20). The region around position 80 of the mature 16S rRNA (#841 in the DNA sequence of Fig. 3) is especially heterogeneous. One difference in this region has affected a HindIII site that is present in *rrnB* but absent in *rrnG*. We have mapped the HindIII sites in all of the *rrn* loci (C.L. Squires, data not shown) and have found this restriction recognition sequence in *rrnA*, *B*, *E* and *H* near position 80 (#841) of the

mature 16S gene. This site is absent in *rrnC*, *D*⁺ and *G*.

It is interesting to examine the differences between the *rrnG* and *rrnB* 16S gene sequences with reference to the 16S rRNA secondary structure model proposed by Noller and Woese (23). That model was derived from *rrnB* sequence data (4). Three of the six differences (#891, #943 and #1010 in Fig. 3) were found in non-base paired regions and three differences lay in areas involved in secondary structure. Two of the latter (#841 and #853) would affect the base paired structure around the sequence corresponding to the HindIII site in *rrnB*; an A:U pair would be changed to a C:U bulge in *rrnG*. At the same time, a G:U pair at the base of the hairpin would be changed to a G:C pair in *rrnG*. The free energies of formation for the two secondary structures would be -4.0 kcal. for *rrnB* versus -0.8 kcal. for *rrnG*, according to the rules of Tinoco *et al* (24). The third *rrnG* heterogeneity would decrease the stability of an *rrnB* structure by changing an A:U pair to a U:U bulge (-8.8 vs. -4.4 kcal.). The actual effect of these heterogeneities on secondary structures or other associations of the 16S rRNA molecules that are derived from *rrnG* is unknown.

DNA structure in *rrn* promoter regions

The *in vitro* (25) and *in vivo* (26) observations that *rrn* promotion is sensitive to inhibitors of DNA gyrase suggest that secondary structure or conformation of DNA plays an important role in the regulation of rRNA expression. Gyrase, through its effect on DNA superhelicity, might be expected to enhance the formation of certain DNA structures. Cruciforms (brought about by same-strand base pairing) and short segments having left-handed DNA conformation have both been demonstrated in negatively supercoiled DNA (27-29). We have examined six *rrn* promoter regions for sequences that could form such structures.

A detailed computer examination of the *rrnG* promoter region sequence for possible same-strand base pairing has revealed one such structure that may have significance (Fig. 5A). The P1 and P2 -10 sequences base pair to form the stem of a cloverleaf-type structure in which 71 percent (89 out of 126) of the nucleotides between the two -10 sequences are paired. Examination of other *rrn* promoter regions showed that only the *rrnA/B* sequence, which is quite similar to that of *rrnG*, supports a similar secondary structure (Fig. 5B). The two structures shown in Figs. 5A and 5B precisely involve the region between P1 and P2, are relatively compact and have the lowest free energies of formation that could be found in the immediate vicinity of either control region. The four regions that differ between the *rrnG* and *rrnB* sequences affect the stability of the two structures. The differences weaken the left and central hairpins of *rrnG* but slightly strengthen the right hairpin. These effects are largely the consequence of the six base difference at positions #525-530. These differences weaken the *rrnG* structure somewhat, without greatly altering the number of nucleotides involved in base pairing.

The most interesting aspect of the structures shown in Fig. 5 is the intimate association of the P1 and P2 -10 sequences. This suggests a concerted involvement of P1 and P2 for the *G* and *B* loci that does not exist for *rrnD*, *E* or *X(H)*. Whether these structures are found under

physiological conditions, or are involved in the expression of the *rrnG* or *rrnB* loci, is unknown. However, the presence of such a secondary structure might be demonstrated *in vitro* by nuclease S1 nicking of the supercoiled DNA (27,28) from M13 clones containing the promoter region [eg. T4(+), Fig. 2].

3311

form left-handed DNA segments, then their transcription start positions would be located in the transition region between the left- and right-handed DNA segments.

Conclusion

Six of the seven *rrn* promoter regions have now been sequenced. With the exception of *rrnA* and *rrnB* all of the promoter sequences show some differences in primary structure. However, they fall into two quite similar classes (*rrnA*, *rrnB* and *rrnG* vs. *rrnD*, *rrnE* and *rrnX(H)*). Whether these two classes indicate a significant difference in the mechanism by which some *rrn* loci are regulated is still an unanswered question. In this paper we have suggested possible structures that would be common to all *rrn* loci (left-handed DNA segments) and other possible structures that could only be formed in the promoter regions of the *rrnA*, *B* and *G* loci (P1-P2 same-strand base pairing). We have also noted that differences may exist in transcription termination of the genes that precede *rrnG* and *rrnB* (ORF-BG and ORF-BB). We are continuing to examine the features described above and to study the possible effects they may have on the regulation of the seven *rrn* loci.

ACKNOWLEDGEMENTS

The authors wish to thank Else Saederup for help with the sequencing. We are also indebted to Louise Clarke and Masayasu Nomura for providing strains, to Dr. Nomura for communicating unpublished results and to Serap Aksoy, Sherman Beychok, Suzanne Li and Alfred Nordheim for helpful suggestions and discussions concerning the manuscript. Computer analysis was supported in part by an NIH grant (RR0042) to the Columbia University Biological Sciences Department Computer Graphics Facility. The remainder of the work was supported by an NIH grant (GM25178) to CLS.

REFERENCES

1. Kenerley, M. E., Morgan, E. A., Post, L., Lindahl, L. and Nomura, M. (1977) *J. Bacteriol.* **132**, 931-949.
2. Zurawsky, G. and Brown, K. (1976) *J. Mol. Biol.* **130**, 83-96.
3. Ellwood, M. and Nomura, M. (1982) *J. Bacteriol.* **149**, 458-468.
4. Brosius, J., Dull, T., Sleeter, D. and Noller, H. (1981) *J. Mol. Biol.* **148**, 107-127.
5. DeBoer, H.A., Gilbert, S.F. and Nomura, M. (1979) *Cell* **17**, 201-209.
6. Young, R. and Steitz, J. (1979) *Cell* **17**, 225-234.
7. Hill, C.W. and Harnish, B.W. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7069-7072.
8. Messing, J., Crea, R. and Seeburg, P. (1981) *Nucl. Acids Res.* **9**, 309-321.
9. Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3642-3646.
10. Clarke, L. and Carbon, J. (1976) *Cell* **9**, 91-99.
11. Spudich, J.A., Korn, V. and Yanofsky, C. (1970) *J. Mol. Biol.* **53**, 49-67.
12. Squires, C., Krainer, A., Barry, G., Shen, W-F. and Squires, C.L. (1981) *Nucl. Acids Res.* **9**, 6827-6840.
13. Staden, R. (1979) *Nuc. Acids Res.* **6**, 2601-2610.
14. Herrman, R., Neugebauer, K., Pirkli, E., Zentgraf, H. and Schaller, H. (1980) *Molec. gen.*

Genet. 177, 231-242.

15. Sanger, F., Coulson, A., Barrell, B., Smith, A. and Roe, B. (1980) *J. Mol. Biol.* 143, 161-178.
16. Van Wezenbech, P., Hulsebos, T. and Schoenmakers, J. (1980) *Gene* 11, 129-148.
17. Rosenberg, M. and Court, D. (1980) *Ann. Rev. Genet.* 13, 319-353.
18. Gilbert, S.F., DeBoer, H.A. and Nomura, M. (1979) *Cell* 17, 211-224.
19. Young, R.A. and Steitz, J. (1978) *Proc. Natl. Acad. Sci. USA* 75, 3593-3597.
20. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.P. (1979) *Eur. J. Biochem.* 100, 399-410.
21. Kingston, R.E., Gutell, R.R., Taylor, A.R. and Chamberlin, M.J. (1981) *J. Mol. Biol.* 146, 433-449.
22. Kingston, R.E. and Chamberlin, M.J. (1981) *Cell* 27, 523-531.
23. Noller, H.F. and Woese, C.R. (1981) *Science* 212, 403-410.
24. Tinoco, 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.
25. Yang, H.-L., Heller, K., Gellert, M. and Zubay, G. (1979) *Proc. Natl. Acad. Sci., USA* 76, 3304-3308.
26. Oostra, B.A., AB, G. and Gruber, M. (1980) *Nucl. Acids Res.* 8, 4235-4246.
27. Panayotatos, N. and Wells, R.D. (1981) *Nature* 289, 466-470.
28. Lilley, D.M.J. (1981) *Nucl. Acids Res.* 9, 1271-1289.
29. Klysik, J., Stirdivant, S.M., Larson, J.E., Hart, P.A. and Wells, R.D. (1981) *Nature* 290, 672-677.
30. Wang, A.H.-J., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Bloom, J.H., van der Marel, G. and Rich, A. (1979) *Nature* 282, 680-686.
31. Arnott, S., Chandrasekaran, R., Birdsall, D.L., Leslie, A.G.W. and Ratliff, R.L. (1980) *Nature* 283, 743-745.
32. Nordheim, A., Pardue, M.L., Lafer, E.M., Moller, A., Stollar, B.D. and Rich, A. (1981) *Nature* 294, 417-422.