Volume 9 Number 12 1981 Nucleic Acids Research

The rRNA operon from Zea mays chloroplasts: nucleotide sequence of 23S rDNA and its homology with E.coli 23S rDNA

K. Edwards and H. Kössel

Institut für Biologie III, Universität Freiburg, Schänzlestrasse 1, 7800 Freiburg, GFR

Received 27 April 1981

ABSTRACT

The nucleotide sequence of 23S rDNA from Zea mays chloroplasts has been determined. Alignment with 23S rDNA from E.coli reveals 71 percent homology when maize 4.5S rDNA is included as an equivalent of the 3' end of E.coli 23S rDNA. Among the conserved sequences are sites for base modification, chloramphenicol sensitivity and ribosomal subunit interaction. A proposal for the base pairs formed between 16S and 23S rRNAs during the 30S/50S subunit interaction is presented.

The alignment of maize 23S rDNA with that of E.coli reveals three small insertion sequences of 25, 65 and 78 base pairs, whereas maize 16S rDNA shows only deletions when compared with the E.coli species.

INTRODUCTION

Determination of the primary structures of ribosomal RNAs is essential to the elucidation of the structure of the ribosome and of the roles played by RNAs in protein synthesis. An important approach to identifying regions of an rRNA species which play a role in ribosome structure or function is the comparison of sequences between rRNAs from different organisms. We have therefore analysed the primary structure of an entire rRNA operon from the chloroplast of Zea mays to allow comparison with the corresponding prokaryotic (1,2,3,4), eukaryotic (5) and mitochondrial (6,7) rRNA species and consequently to provide a quantitative parameter for determining phylogenetic relationships. Secondly, such a comparison can help to confirm or exclude predictions of secondary structures for rRNA species (8,9) and this "phylogenetic criteria" has already been useful in the formulation of a 16S rRNA secondary structure (10,11). In addition, de-

termination of the primary structure of a chloroplast rRNA operon provides a basis for studying the various steps of its expression. Since the expression of all protein coding genes of the chloroplast genome is dependent on organelle specific ribosomes, the structure and regulation of chloroplast rRNA operons play a key role in the biogenesis of the organelle.

DNA from Zea mays chloroplasts is a circular molecule of about 135 kilobase pairs (kbp) on which two rRNA coding regions are positioned in opposite orientation within two 22 kbp inverted repeats (12). Each inverted repeat includes within a 12 kbp EcoRI fragment one copy of a 16S rRNA, 23S rRNA, 4.5S and 5S rRNA gene and a 2.4 kbp spacer region between the 16S and 23S rRNA genes. One such 12 kbp fragment has been linked to the plasmid vector pMB9 within the E.coli clone pZmc134 (12). This clone was used for the isolation and sequencing of 23S rRNA coding DNA fragments.

We wish to present the complete primary structure of the 23S rRNA gene from Zea mays chloroplasts. Together with the sequences for the 16S rDNA (9), the 16S - 23S spacer rDNA (Koch, W., Edwards, K. and Kössel, H., submitted for publication), a 600 base pair leader region proximal to the 16S rRNA gene (13) and the 4.5S/5S sequence distal to the 23S rRNA gene (Bedbrook, J.R. and Dyer, T., submitted for publication), this completes the primary structure of one entire rRNA operon from Zea mays chloroplasts.

MATERIALS AND METHODS

Materials

Restriction endonucleases were purchased from New England Biolabs, Bethesda Research Laboratories and Boehringer-Mannheim. Polynucleotide kinase was from New England Biolabs, DNA polymerase was a gift of Dr. L. Loeb, Institute for Cancer Research, Fox Chase, Philadelphia, and bacterial alkaline phosphatase was obtained from Worthington. [α-^{32}P] -dNTPs (400 Ci/nmol) and [γ-^{32}P] -ATP (> 5000 Ci/nmol) were purchased from the Radiochemical Centre, Amersham, England. The plasmid clone pZmc134 (12)

was a gift of Dr. J. Bedbrook.

Preparation of DNA and mapping with restriction enzymes

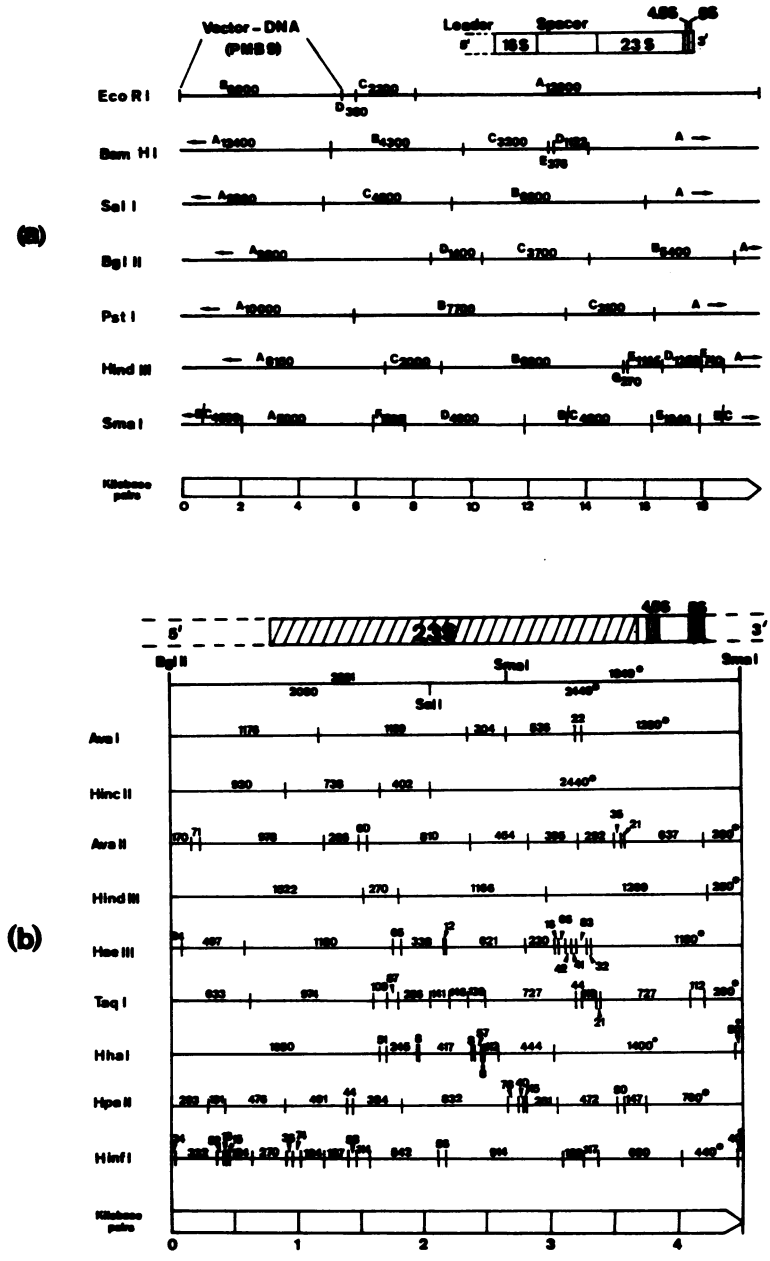
Growth of the E.coli clone pZmc134 and preparation of its plasmid DNA were as described previously (12). Restriction enzyme digests were carried out as specified by the manufacturers. The restriction map of the plasmid pZmc134 for the enzymes EcoRI, BamHI, SalI, BglII, PstI, HindIII and SmaI is shown in Fig. 1a. The fragments $BglII \cdot B_{5400}$, $BamHI \cdot A_{12400}$ or $SmaI \cdot E_{1840}$ were purified by electrophoresis on 1% agarose gels and further digested with restriction endonucleases to give the fine map shown in Fig. 1b and to yield fragments suitable for DNA sequencing. Fragments smaller than 40 base pairs (bp) were not detected on the mapping gels used and hence the presence of such fragments in Fig. 1b was determined from the subsequent DNA sequencing alone.

Labelling and sequencing of DNA fragments

DNA fragments were either 3'-end labelled with DNA polymerase and 20 µCi of a suitable $[\alpha\text{-}^{32}P]$-dNTP (14,15) or 5'-end labelled with polynucleotide kinase and 75 µCi $[\gamma\text{-}^{32}P]$-ATP after treatment with alkaline phosphatase and phenol extraction (15).

Labelled fragments were separated on 5% or 7.5% polyacrylamide gels, extracted (15) and either cut with a second restriction enzyme or strand separated on 7.5% (fragments < 150 bp) or 6% (fragments > 150 bp) polyacrylamide gels as described (15). Following extraction from the gels, all fragments or single strands were cleaned using the hexadecyltrimethylammonium - bromide-butanol extraction method (16) before being subjected to the hydrazine/dimethyl-sulphate sequencing methode (15).

The sequencing strategy is outlined in Fig. 1c. For most of the fragments a series of gels were used which gave overlapping sequences (not individually indicated in Fig. 1c), allowing multiple reading of more than 90% of all positions.


(a)
Vector - DNA
(PMB9)
Leader
Spacer
16 S
23 S
5'
3'
Eco R I
Bam H I
Sal I
Bgl II
Pst I
Hind III
Sma I
Kilobase pairs
(b)
Bgl II
Sma I
Sal I
Ava I
Hinc II
Ava II
Hind III
Hae III
Taq I
Hha I
Hpa II
Hinf I
Kilobase pairs

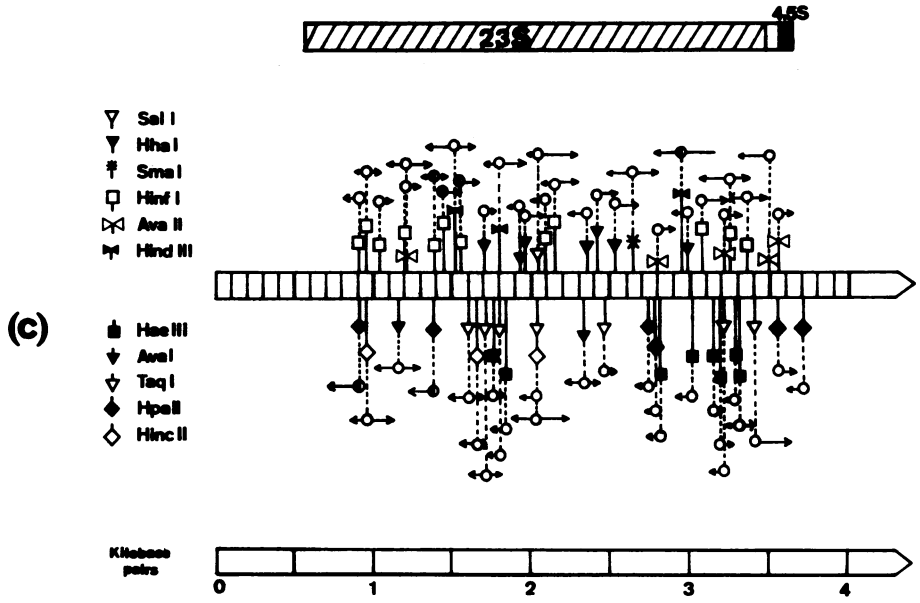
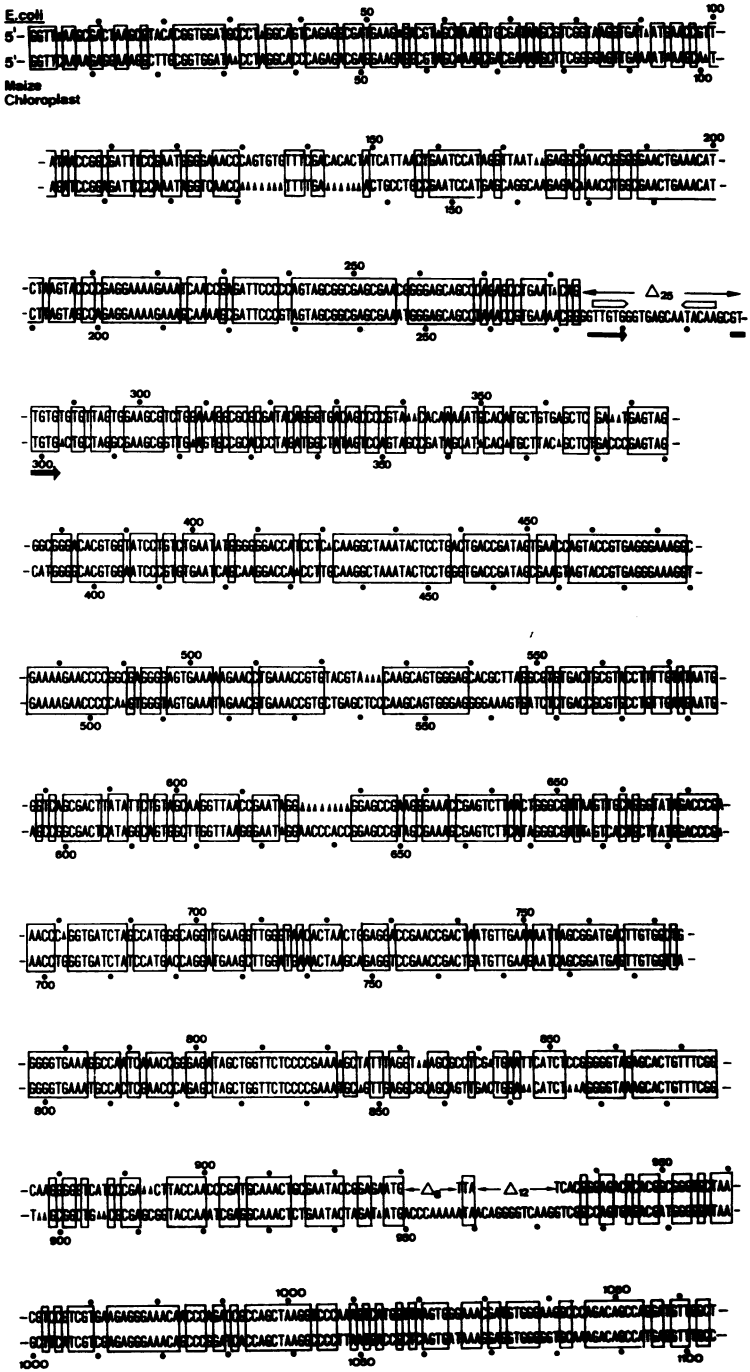


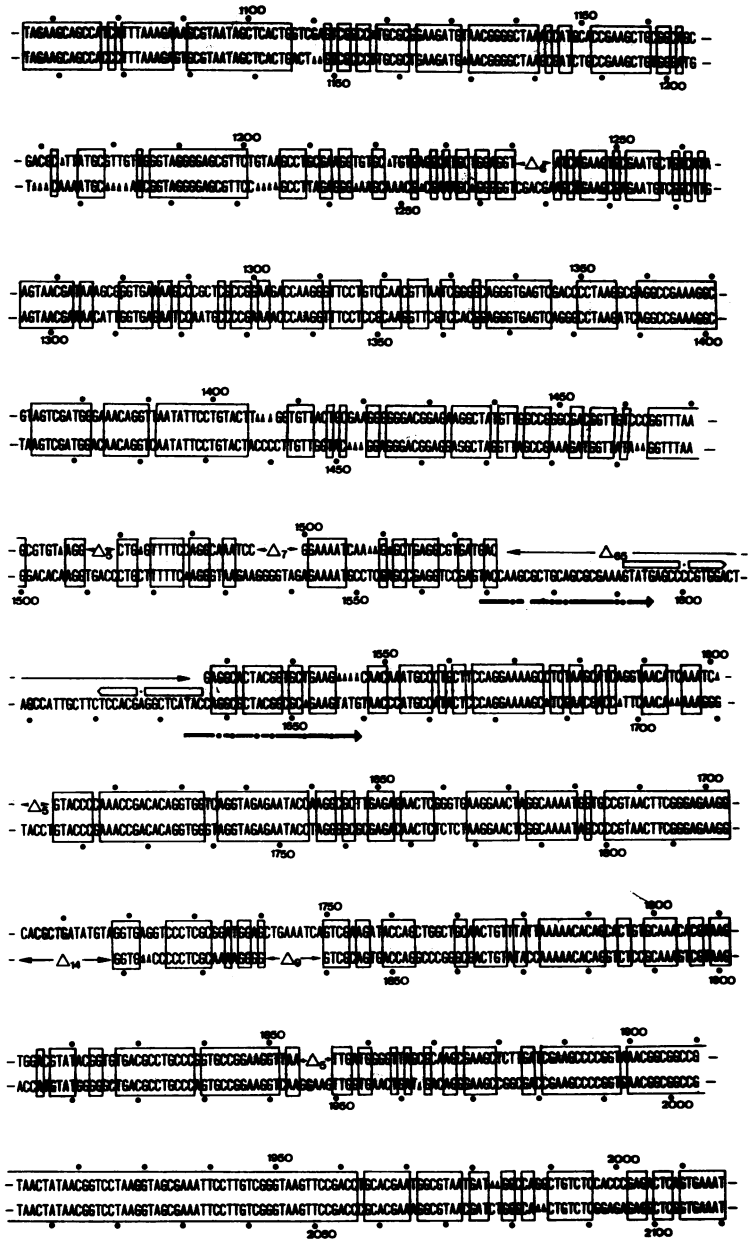
Fig. 1. Physical mapping and use of restriction fragments for sequence analysis of 23S rDNA from Zea mays chloroplasts. In a the restriction map for hexamer recognizing restriction endonucleases is shown relative to the position of cloned rDNA from maize chloroplasts (fragment EcoRI·A). b represents the fine mapping of the fragments BglII·B/SmaI$_{2651}$, SmaI·E$_{1840}$ and HindIII·E$_{1166}$. Asterisks denote that the fragment size is approximate. In c the individual fragment ends selected for terminal labelling and sequence analysis of the respective fragments are indicated by open circles (5'-terminally labelled) or half-filled circles (both 5'- and 3'-terminally labelled, in separate experiments).

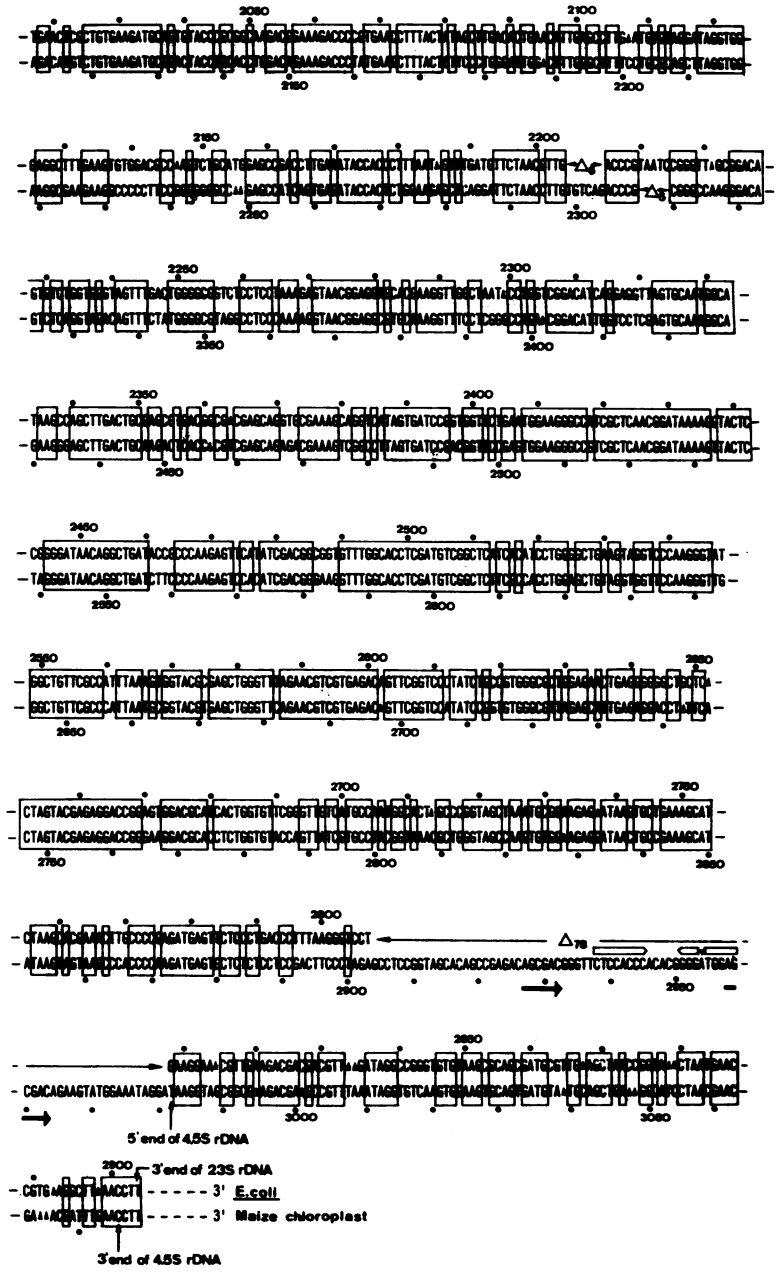
RESULTS AND DISCUSSION

Sequence of maize chloroplast 23S rRNA gene and its homology with that of E.coli

The RNA-like strand of the entire 23S rRNA gene from maize chloroplast is depicted in Fig. 2 (lower row of each line). It represents the first sequence of a complete chloroplast 23S rRNA gene, although partial sequences from the intron flanking regions of Chlamydomonas reinhardii chloroplast 23S rDNA (17) and from the 5' end of Euglena gracilis chloroplast 23S rDNA (18,19) have been published. We have determined the maize sequence up to position 2992 (Fig. 2),


E.coli
5'-
5'-
Maize
Chloroplast




5' end of 4.5S rDNA
3' end of 23S rDNA
3' E.coli
3' Maize chloroplast
3' end of 4.5S rDNA

giving an 84 bp overlap with the sequence of the 4.5S/5S rDNA region determined by Bedbrook and Dyer (submitted for publication).

Extensive homology is observed between the maize chloroplast 23S rDNA sequence and the 23S rDNA (4) or 23S rRNA (3) sequences from E.coli (Fig. 2). As has already been noted (20,21), chloroplast 4.5S rRNA shows strong sequence homology to the 3' end of 23S rRNA from E.coli, and we have therefore included the maize chloroplast 4.5S rDNA sequence in the alignment shown at the 3' ends of the sequences in Fig. 2. A total of 2059 positions (71%) of the 2904 residues from E.coli 23S rDNA are identical in maize when the alignment depicted in Fig. 2 is used for comparison of the two sequences. The longest stretch of consecutive homologous nucleotides (63) is located between positions 1994 and 2056 of maize 23S rDNA and many shorter homologous sequences are scattered throughout the two genes.

The exact 5' and 3' termini of the 23S rRNA gene are unknown since the maize 23S rRNA has not been sequenced. The 5' terminus shown in Fig. 2 has, therefore, been tentatively assigned on the basis of homology with E.coli 23S rDNA. The 3' terminus must lie proximal to the beginning of the 4.5S rDNA (Fig. 2), probably in the region between bases 2880 and 2981 (maize numbering), where no homology to the E.coli sequence can be detected.

Comparison with large subunit RNAs from species other than E.coli

The 5' terminal regions of E.coli and maize chloroplast 23S rDNAs show a high degree of homology to that of 23S rDNA

Fig. 2. The nucleotide sequence (RNA-like strand) of the 23S rDNA from Zea mays chloroplasts is shown on the lower row of each line. The RNA-like strand of E.coli 23S rDNA (4) is depicted in the upper row for comparison. The 4.5S rDNA sequence from maize (Bedbrook, J.R. and Dyer, T., submitted for publication) is also shown. Homologous sequences are marked by boxes. Positions deleted as compared to the other DNA are shown as Δ. Open arrows (⇨) represent inverted repeats within the 25 bp, 65 bp and 78 bp inserts. Filled arrows (➡) represent tandem repeats in the same regions.

from Euglena gracilis chloroplasts. Intron flanking sequences in the 25S rDNA from Tetrahymena pigmentosa (22), the 23S rDNA from Chlamydomonas reinhardii chloroplasts (17) and the 21S rDNA from yeast mitochondria (24) are very similar to the corresponding sequences from E.coli and maize chloroplast. However, no intron or small insertion sequences are observed between the corresponding positions 2019/2020, 2546/2547 and 2689/2690 in maize. The intron flanking sequences from Tetrahymena (22) are also well conserved in Xenopus laevis 28S rRNA (23), although no intron is present in this region in Xenopus.

The complete sequences of large subunit rDNAs (coding for 16S rRNA species) from human (6) and mouse (7) mitochondria are available for comparison. The alignment with maize 23S rDNA reveals limited homology and closely resembles the homologies observed between mouse mitochondrial 16S rDNA and E.coli 23S rDNA (7) (K. Edwards, in preparation).

Location of specific regions in 23S rRNA

All sites known to be post-transcriptionally modified in E.coli 23S rRNA (3,4) have been conserved in the maize 23S rDNA sequence, and these sites are all found in extended stretches of homology.

Two chloramphenicol-resistant mutants in yeast mitochondrial 21S rRNA have been analyzed and found to contain single base substitutions (24) at sites corresponding to positions 2447 and 2503 of E.coli 23S rDNA. Both of these bases are conserved in the maize chloroplast species (corresponding to positions 2543 and 2599) and are located within stretches of 16 and 23 identical nucleotides, respectively, when compared to the E.coli 23S rDNA. Comparison of the maize and E.coli sequences surrounding these two bases with the corresponding sequences in yeast 21S rRNA shows that the two bases are located in stretches of 11 and 13 nucleotides which are identical between the three species. As previously noted (4), this region of the 23S rRNA may be involved in the peptidyl transferase function and the strong conservation observed in this region is most likely due to structural constraints necessary to preserve this function.

Complementarity of 23S rRNA to 5S rRNA and to $tRNA_F^{Met}$

In E.coli 23S rRNA the nucleotides at positions 143-154 were found to be complementary to nucleotides 72-83 of 5S RNA (4). However, there is no homology between maize and E.coli 23S rDNA in this region, and no complementarity between this region of maize 23S rDNA and maize 5S rDNA (Bedbrook, J.R. and Dyer, T., submitted for publication). A region complementary to nucleotides 68-75 of maize 5S rDNA has been found at positions 1846-1855 of maize 23S rDNA and fragment/cross-linking data also indicate an interaction between the corresponding positions of the 5S and 23S rRNAs in E.coli (Zwieb, C., Glotz, C. and Brimacombe, R., personal communication).

In E.coli 23S rRNA positions 1984-2001 are complementary to $tRNA_F^{Met}$ (25). This region is only partly conserved in the maize sequence and allows only 7 base pairs (divided into 4 plus 3 by one mismatch) with $tRNA_F^{Met}$ from bean or spinach chloroplasts (26,27). This limited conservation can be interpreted as follows:

(a) this interaction has no structural or functional significance in either E.coli or maize chloroplasts;

(b) this interaction takes place in E.coli only;

(c) this interaction is functionally significant in both E.coli and maize chloroplasts but is limited in both, or in maize only, to 7 base pairs.

Possible sites of interaction between the 30S and 50S ribosomal subunits

Interaction between the 30S and 50S ribosomal subunits may be mediated either by protein-protein interaction, protein-RNA interaction or RNA-RNA interaction. In the latter, the regions of the 16S and 23S rRNAs which are involved in the subunit interaction may, in their non-interacting states, be either base paired with other regions of the same RNA or exist in non-base paired areas of the molecule (e.g. at tops or bottoms of stems in the secondary structure). If they exist in non-base paired regions, G residues present in the sequence should be kethoxal reactive and reactivity should be suppressed in 70S compared to either 30S or 50S subunits

alone. In addition, reassociation of 30S with 50S subunits should be lost following kethoxal modification of G residues in regions of either 16S or 23S rRNA which are involved in subunit interaction.

Using these two criteria, Noller et al. have listed seven fragments from 16S rRNA (28,29) and three from 23S rRNA (30) which are potentially involved in 30S/50S subunit interaction and have proposed two base pairing interactions between 16S and 23S rRNA (29). We have also used this fragment list to propose four additional possible 16S-23S rRNA interaction sites (Table 1) to which we have applied the following criteria:

(i) A minimum of three base pairs between 16S and 23S rRNA interacting sites.

(ii) Both the kethoxal reactive G residues in the 16S and in the 23S rRNAs are base paired.

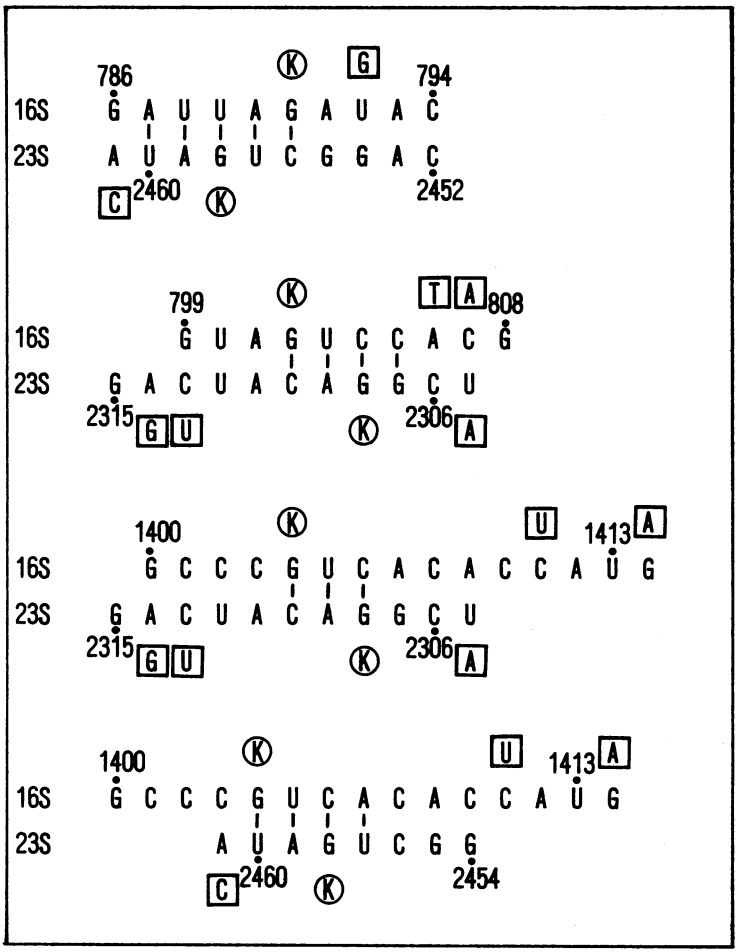
(iii) The base pairing between the 16S and 23S rRNAs in E.coli is conserved either by identical sequences or by compensating base changes in maize chloroplasts.

We regard three base pairs as being sufficient for the interaction, since the stabilizing energy would come not from base pairing alone but also from the stacking energy of neighbouring bases, in a manner similar to the pairing observed between tRNA molecules containing complementary anticodons (31). In addition, long stretches of base pairing may be unfavourable because, due to the necessary helical turns, they would confer a severe topological constraint on the interaction. This would not be the case if the interaction involves only three to five base pairs.

The two interactions between the 16S and 23S rRNAs previously proposed (29) do not fulfill either criteria (ii) or (iii) and have therefore not been considered in Table 1. The interactions proposed in Table 1 are further supported by strong conservation of the 16S sequences throughout prokaryotes (32) and also in yeast 18S RNA (5).

Several of the sequences listed in Table 1 are involved in more than one base pairing interaction. We consider it possible that a given sequence may be involved in interchain

Table 1: Possible base pairing interactions between E.coli 16S and 23S rRNAs involved in 30S/50S subunit interaction.



Base substitutions in maize are shown in squares. (K) designates kethoxal reactive G residues. Numbering is from E.coli 16S rRNA (1,2) and 23S rRNA (3,4).

switches between 16S and 23S rRNAs in a way similar to that already proposed for intrachain switches of base pairs within the 16S secondary structure (33).

Presence of insertion sequences

The alignment illustrated in Fig. 2 contains the following three insertions in the maize chloroplast 23S rDNA se-

quence:
(a) a 25 bp insert at positions 274-298,
(b) a 65 bp insert between bases 1572 and 1636, and
(c) a 78 bp insert at positions 2904-2981.

Each of the inserts shows some structural resemblance to prokaryotic insertion elements which are characterized by the presence of inverted repeats and flanking tandem repeats (for review see 34). The inverted repeats present in the inserts are shown as open arrows above the maize sequence in Fig. 2, while the tandem repeats are shown as filled arrows below the sequence. The 25 bp insert contains a 5 bp inverted repeat which enables the stem structure shown in Fig. 3a to be drawn and is flanked by a 6 bp tandem repeat. The inverted repeat present in the 65 bp insert gives rise to the stem structure depicted in Fig. 3b. The 65 bp insert is flanked by a long, interrupted tandem repeat. Similarly, tandem repeats are present in the 78 bp insert and an inverted repeat containing one mismatch leads to the structure shown in Fig. 3c. It is interesting to note that a mini-insert of 66 bp which also shows a terminal repeat at its ends and a palindromic sequence in its middle region is present in yeast 21S rDNA (24).

This structural resemblance to bacterial insertion elements (34) leads to the speculation that the insertion sequences observed in the maize 23S rDNA are either small IS-like elements which are still functional or represent remnants of former IS-elements or larger intron sequences. The latter proposal is supported by the detection of intron sequences in the large subunit rDNAs of many species (for survey see ref. 23). In addition, intron sequences have been found within the two tRNA genes located in the 16S-23S spacer rDNA of maize chloroplasts (35). It remains to be seen whether the RNA sequences corresponding to the insertion sequences are present in the 23S rRNA of mature ribosomes or if they are excised during rRNA maturation.

Since the 4.5S rRNA is known to exist in the ribosomes of higher plant chloroplasts as a species separate from the 23S rRNA (36,37,38) it follows that a primary transcript

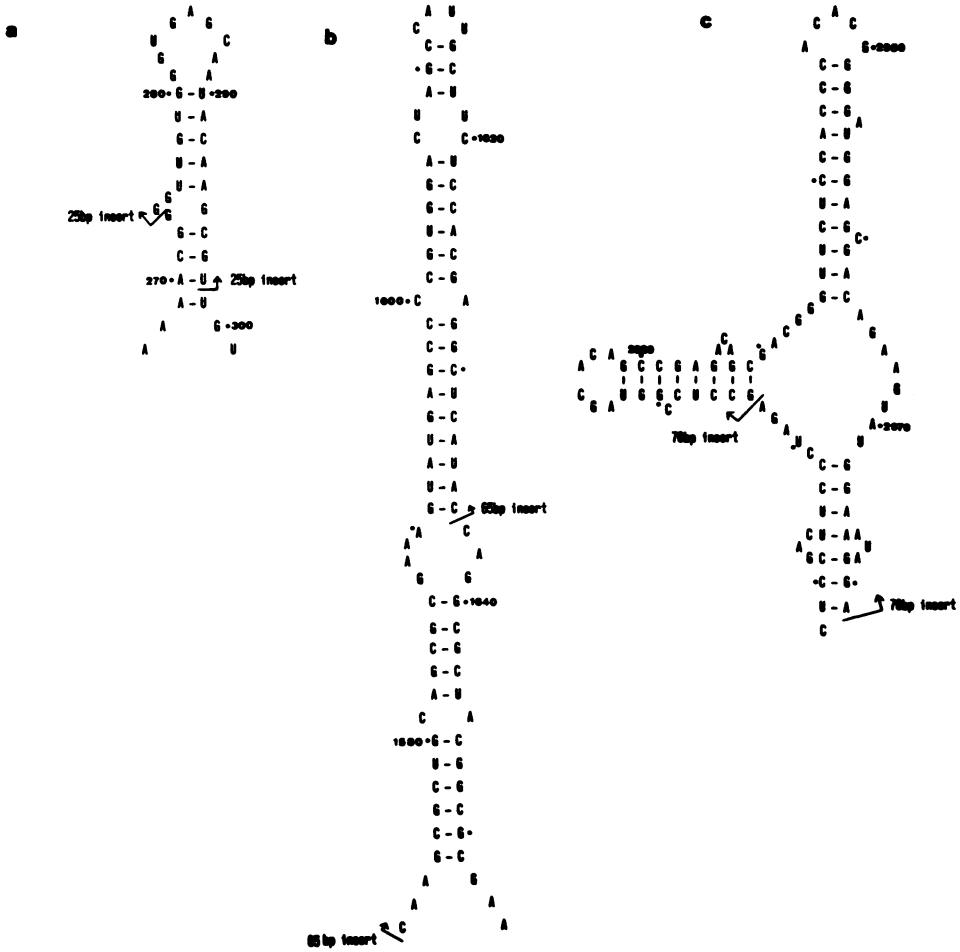


Fig. 3. Possible secondary structures formed by the a 25 base pair insert, b 65 base pair insert, c 78 base pair insert and adjacent nucleotides in maize 23S rRNA.

containing both 4.5S and 23S rRNA must be cut between the regions containing the 23S rRNA and the 4.5S rRNA. It appears likely that the signal structure(s) necessary for this cut is encoded in, and has been introduced during evolution by, the 78 bp insert. In particular, the stem structure depicted in Fig. 3c may represent a signal structure for an RNase III like processing enzyme.

ACKNOWLEDGEMENTS

We thank Dr. J. Bedbrook for kindly providing us with the clone pZmc134, Mrs. E. Schiefermayr for technical assistance and Drs. R. Hertel and P. Langridge for helpful criticism of the manuscript. We would also like to thank Drs. J. Bedbrook and T. Dyer for kindly providing us with the 4.5S rDNA sequence prior to publication. This work was supported by grants from the Deutsche Forschungsgemeinschaft (SFB 46) to H.K.

REFERENCES

1. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc.Natl.Acad.Sci. USA 75, 4801-4805
2. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.P. (1978) FEBS Letters 94, 152-156
3. Branlant, C., Krol, A., Machatt, M.A. and Ebel, J.P. (1979) FEBS Letters 107, 177-181
4. Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc.Natl.Acad.Sci. USA 77, 201-204
5. Rubtsov, P.M., Musakhanov, M.M., Batchikova, N.V., Skryabin, K.G. and Bayev, A.A. (1979) Dokl.Adad.Nauk SSR 248, 760-762
6. Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980) Nature 286, 460-467
7. Van Etten, R.A., Walberg, M.W. and Clayton, D.A. (1980) Cell 22, 157-170
8. Schwarz, Zs. and Kössel, H. (1979) Nature 279, 520-522
9. Schwarz, Zs. and Kössel, H. (1980) Nature 283, 739-742
10. Glotz, C. and Brimacombe, R. (1980) Nucleic Acids Res. 8, 2377-2395
11. Woese, C.R., Magrum, L.J., Gupta, R., Siegel, R.B., Stahl, D.A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J.J. and Noller, H.F. (1980) Nucleic Acids Res. 8, 2275-2293
12. Bedbrook, J.R., Kolodner, R. and Bogorad, L. (1977) Cell 11, 739-749
13. Schwarz, Zs., Kössel, H., Schwarz, E. and Bogorad, L. (1981) Proc.Natl.Acad.Sci. USA, in press
14. Schwarz, E., Scherer, G., Hobom, G. and Kössel, H. (1978) Nature 272, 410-414
15. Maxam, A. and Gilbert, W. (1980) in Methods of Enzymology 65, Grossman, L. and Moldave, K., Eds., pp. 499-560, Academic Press, New York
16. Langridge, J., Langridge, P. and Bergquist, P.L. (1980) Anal.Biochem. 103, 264-271
17. Allet, B. and Rochaix, J.-D. (1979) Cell 18, 55-60
18. Graf, L., Kössel, H. and Stutz, E. (1980) Nature 286, 908-910
19. Orozco, E.M., Rushlow, K.E., Dodd, J.R. and Hallick, R.B. (1980) J.Biol.Chem. 255, 10997-11001

20. MacKay, R.M. (1981) FEBS Letters 123, 17-18
21. Edwards, K., Bedbrook, J., Dyer, T. and Kössel, H. (1981) Biochemistry International, in press
22. Wild, M.A. and Sommer, R. (1980) Nature 283, 693-694
23. Gourse, R.L. and Gerbi, S.A. (1980) Nucleic Acids Res. 8, 3623-3637
24. Dujon, B. (1980) Cell 20, 185-197
25. Dahlberg, J.E., Kintner, C. and Lund, E. (1978) Proc. Natl.Acad.Sci. USA 75, 1071-1075
26. Canaday, J., Guillemaut, P. and Weil, J.H. (1980) Nucleic Acids Res. 8, 999-1008
27. Kashdan, M., Pirtle, R., Pirtle, I., Calagan, J., Vreman, H. and Dudock, B. (1980) Fed.Proc. 39, 2022 (Abstract 2192)
28. Herr, W. and Noller, H.F. (1978) Biochem. 17, 307-315
29. Herr, W., Chapman, N.M. and Noller, H.F. (1979) J.Mol.Biol. 130, 433-449
30. Herr, W. and Noller, H.F. (1979) J.Mol.Biol. 130, 421-432
31. Grosjean, H., Söll, D.G. and Crothers, D.M. (1976) J.Mol.Biol. 103, 499-519
32. Woese, C.R., Fox, G.E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B.J. and Stahl, D. (1975) Nature 254, 83-86
33. Brimacombe, R. (1980) Biochemistry International 1, 162-171
34. Calos, M.P. and Miller, J.H. (1980) Cell 20, 579-595
35. Koch, W., Edwards, K. and Kössel, H. (1980) 13th FEBS Meeting, Jerusalem, Abstract S3-P32, p. 61
36. Dyer, T.A. and Bowman, C.M. (1976) in Genetics and Biogenesis of Chloroplasts and Mitochondria, Bücher, Th., Neupert, W., Sebald, W. and Werner, S., Eds., pp. 645-651, North-Holland Publishing Co., Amsterdam and New York
37. Dyer, T.A., Bowman, C.M. and Payne, P.I. (1977) in Nucleic Acids and Protein Synthesis in Plants, Bogorad, L. and Weil, J.H., Eds., pp. 121-133, Plenum Press, New York and London
38. Whitfeld, P.R., Leaver, C.J., Bottomley, W. and Atchison, B.A. (1978) Biochem.J. 175, 1103-1112