Volume 10 Number 23 1982 Nucleic Acids Research

Sequence of the chloroplast 16S rRNA gene and its surrounding regions of *Chlamydomonas reinhardii*

M.Dron, M.Rahire and J.-D.Rochaix+

Département de Biologie moléculaire, Université de Genève, 30, quai Ernest-Ansermet, 1211 Genève 4, Switzerland

Received 20 September 1982; Revised and Accepted 8 November 1982

ABSTRACT

The sequence of a 2 kb DNA fragment containing the chloroplast 16S ribosomal RNA gene from Chlamydomonas reinhardii and its flanking regions has been determined. The algal 16S rRNA sequence (1475 nucleotides) and secondary structure are highly related to those found in bacteria and in the chloroplasts of higher plants. In contrast, the flanking regions are very different. In C. reinhardii the 16S rRNA gene is surrounded by AT rich segments of about 180 bases, which are followed by a long stretch of complementary bases separated from each other by 1833 nucleotides. It is likely that these structures play an important role in the folding and processing of the precursor of 16S rRNA. The primary and secondary structures of the binding sites of two ribosomal proteins in the 16SrRNAs of E. coli and C. reinhardii are considerably related.

INTRODUCTION

Very similar secondary structure models have recently been proposed for E. coli 16S ribosomal RNA (1,2,3). They are based on several independent approaches such as comparison of extensive sequence data of RNA from the small ribosomal subunit in bacteria, chloroplasts, mitochondria and eukaryotes, studies of single strand specific chemical modification of nucleotides, nuclease sensitivity of RNA structure and isolation of complexes of base-paired RNA fragments. Using this model, Stiegler et al. (4) have proposed a general secondary structure for all small ribosomal subunit RNAs in prokaryotes and eukaryotes in which many features of the primary and secondary structure proposed for E. coli 16S rRNA are conserved. Structures of this sort very likely play an important functional role in the ribosome.

The analysis of ribonuclease T1 fingerprints of chloroplast 16S rRNAs has revealed that they share many common

0305-1048/82/1023-7609 $2.00/0

features with prokaryotic rRNAs (5,6). Two higher plant chloroplast 16S rRNA sequences have been established recently in maize and tobacco (7,8). These sequences are highly homologous with each other (95%) and they share 75% of their nucleotides with E. coli 16S rRNA.

The chloroplast 23S rRNA gene region of C. reinhardii presents unique features. The 5' end of the 23SrRNA gene has been split into the small 7S and 3S rRNA genes (9) and a 0.87 kb intron is present near the 3' end of the gene (10). In addition, sequences homologous to the higher plant chloroplast 4.5S rRNA constitute the 3'end of the C. reinhardii 23S RNA (unpublished observations). These observations prompted us to examine whether the 16S rRNA gene of C. reinhardii also possesses distinctive properties.

Here we report the complete sequence of a 1968bp chloroplast DNA segment which contains the entire 16S rRNA gene of C. reinhardii and its flanking regions. This 16S rRNA sequence which is strongly related to the corresponding sequences in E. coli and in the chloroplasts of higher plants is compatible, with minor exceptions, with the general model proposed for 16S rRNA (4). The surrounding regions of the gene are complementary to each other and they are separated from the two ends of the gene by short AT rich regions.

MATERIALS AND METHODS

(a) Plasmids

The chloroplast ribosomal BamHI fragment Ba4 (10) which contains the entire 16S rRNA gene, the 16S-23S spacer and the 7S, 3S and portions of the 23S rRNA genes was used for this work (cf. fig. 1). The fragments BR 1.3 and HR 1.14 were subcloned into the plasmid pBR322 (11) and used for fine structure mapping and sequencing. Plasmids were prepared as described (12).

(b) Restriction enzyme mapping

Restriction enzymes were purchased from New England Biolabs and used as suggested by the manufacturer. The restriction site mapping was performed by partial digestions of end labelled DNA fragments (13) and the maps were confirmed by

complete digestion of the DNA fragments.

(c) DNA sequencing

DNA fragments were end labelled at their 5' and 3' ends as described (14). End labelled fragments were either cut asymmetrically and fractionated by polyacrylamide gel electrophoresis or strand separated (14). DNA sequencing was carried out by the chemical cleavage method of Maxam and Gilbert (14). The reaction products were separated on 8% and 20% thin polyacrylamide sequencing gels (14). The DNA sequence analysis was performed on a Hewlett Packard computer, model 9845.

(d) S1 nuclease mapping

The 1.4 kb BamHI-TaqI fragment (on the left end of Ba4 in fig. 1) was labelled at its 5'ends and strand separated on a 4% polyacrylamide gel. The two strands were hybridized separately with cellular RNA prepared as described (10) and the hybrids were digested with S1 nuclease (15). The S1 resistant DNA products were sized on 5% sequencing gels using the partial cleavage products of the sequencing reaction as size standards.

RESULTS AND DISCUSSION

The chloroplast 16S rRNA gene of C. reinhardii has been mapped previously by electron microscopy (10). The 3' end of the gene was positioned very near the right EcoRI site of fragment RO7 in fig. 1. This figure shows the restriction map of the 16S rRNA gene region and the sequencing strategy used. The continuity of the sequence on both sides of the EcoRI site near the middle of the 16S rRNA gene was not demonstrated by overlap sequencing since we know that fragments BR 1.3 and RO7 (fig. 1) are contiguous on the chloroplast DNA map (unpublished observations).

The 5' end of the mature 16S rRNA was determined by S1 nuclease mapping. As shown in fig. 2, the end of the protected fragment is 64 ± 2 nucleotides away from the left most TaqI site of fig. 1. Since the larger size X and Y bands do not appear reproducibly, we do not know whether they are due to an artefact or whether they correspond to precursors of 16S rRNA. The 3' end of the 16S rRNA was tentatively identified from its

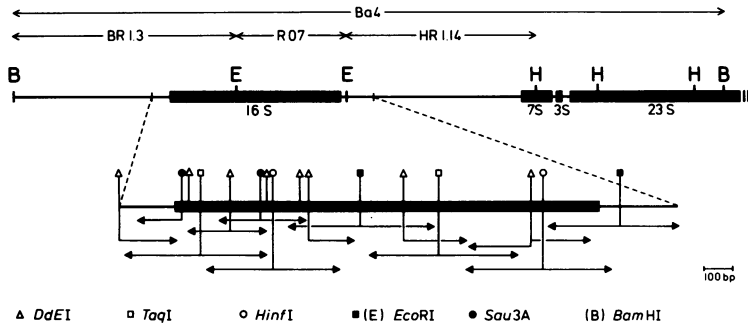

Figure 1.
Restriction endonuclease map and strategy used to sequence the chloroplast 16S rRNA gene region of C. reinhardii. Transcription of the 16S rRNA gene proceeds from the left to the right. Restriction sites linked to vertical lines were either 5'- or 3'-end labelled. The direction and extent of each sequencing run is indicated by an arrow.

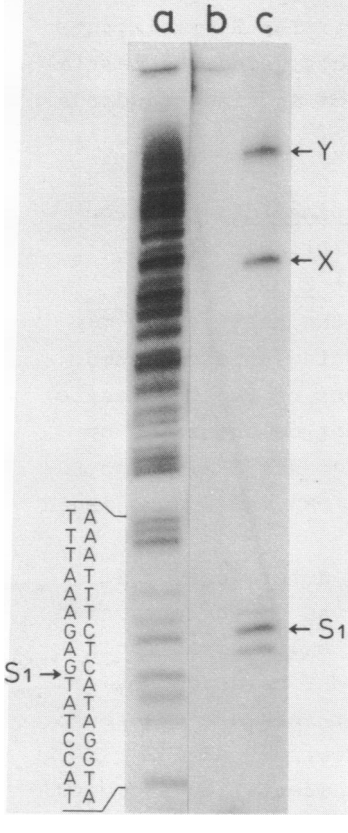

Figure 2.
Location of the 5' end of the chloroplast 16S rRNA gene of C. reinhardii. (a) A≻ C cleavage tract of 5' end labelled fragment TaqI - BamHI (cf. fig. 1. and Materials and Methods). The S1 nuclease protected DNA fragments obtained after hybridization of the Taq -BamHI fragment with yeast RNA (b) or C. reinhardii RNA (c) were electrophoresed in parallel with the partial cleavage products of (a). Since the fragments produced by the sequencing reactions are known to migrate 1 1/2 nucleotide faster than the corresponding fragments obtained by S1 nuclease digestion (22), the 5' end of the 16SrRNA gene is ATCC.

remarkable sequence relatedness to the homologous E. coli sequence. Fig. 3 compares the 16S rRNA gene sequences of C. reinhardii and tobacco (8). Excluding the small deletions and insertions, the two sequences are 85% homologous. The sequence homology with E. coli is 76% (16). The C. reinhardii chloroplast 16S rRNA consists of 1475 ± 2 nucleotides, 11, 16 and 67 bases smaller than the corresponding RNAs from tobacco (8), maize (7) and E. coli (16).

While the chloroplast 16S rRNA coding sequence is highly related to those of higher plants this is no longer true for the gene flanking regions, although these flanking regions are considerably conserved in tobacco, maize and spinach chloroplasts (8; 17; Briat, Dron, Loiseaux and Mache, unpublished results).

It is well documented that the surrounding sequences of the E. coli 16S rRNA gene contain complementary sequences which form an RNAse III cleavage site (18). Similar structures can also be formed in maize (17) and tobacco (8) chloroplasts, while in Euglena the chloroplast 16S rRNA gene is flanked by homologous sequences (19). In C. reinhardii the surrounding regions also contain complementary sequences over at least 44 bases with a single bulge loop of 2 bases (figs 3 and 4). The base complementarity starts about 180 nucleotides from both ends of the 16S rRNA gene (fig. 4); the 5' and 3' intervening segments have an AT content of 75 end 72.5% respectively, considerably higher than for the coding region (fig. 5). The role of these sequences in the processing of the ribosomal precursor is not clear.

In E. coli the 3' end of the 16S rRNA plays an important role in the initiation of protein synthesis since it is involved in the binding of the mRNA in the 5' upstream region of the initiation codon (20). In C. reinhardii chloroplasts base complementary exists, too, between the 3'end of the 16S rRNA and a region of chloroplast mRNAs located several nucleotides upstream of the initiation codon (21).

The sequence of the chloroplast 16S rRNA of C. reinhardii is compatible with the general secondary model proposed by Stiegler et al. (fig. 6 of ref. 4) with minor modifications.

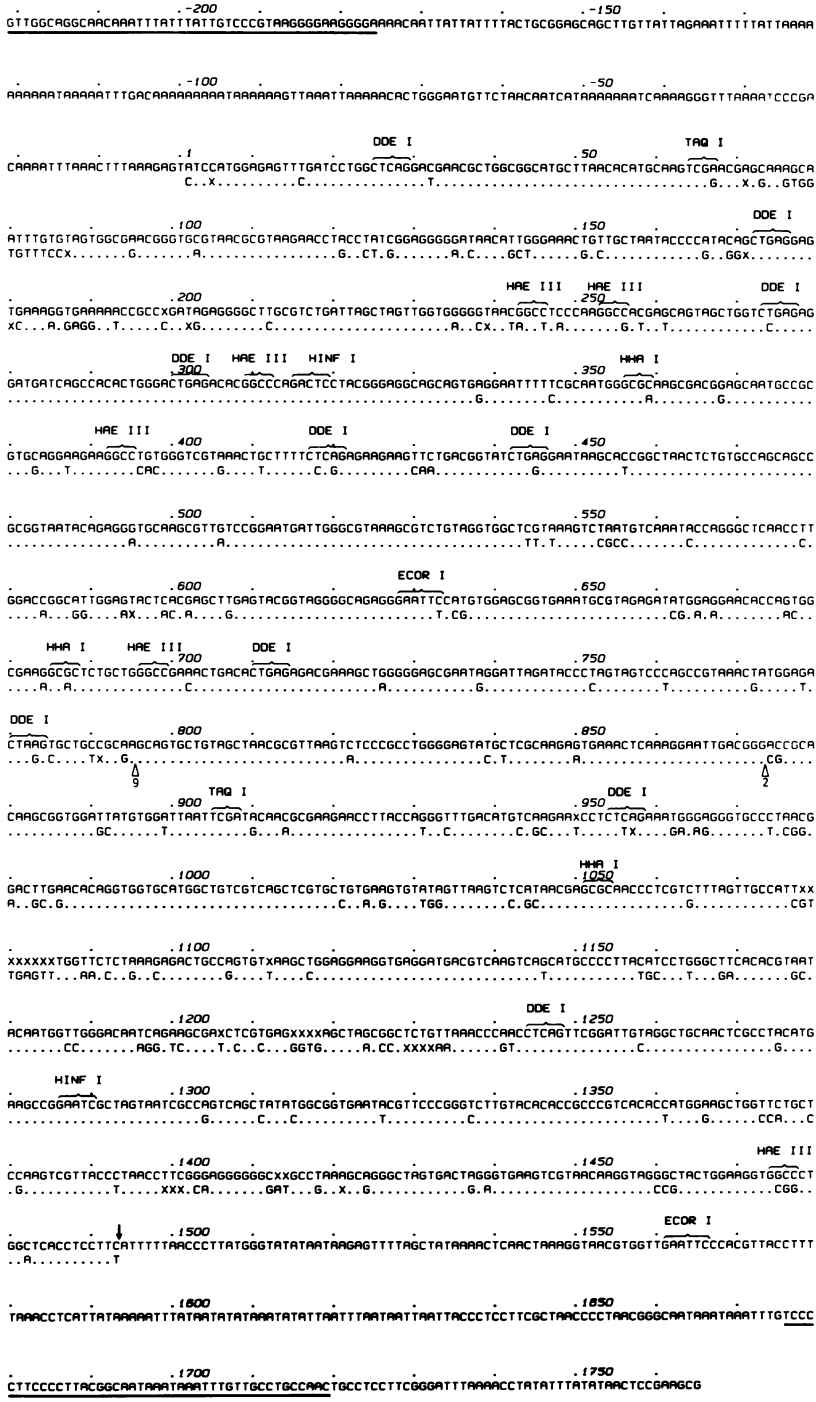

Table I
Deletions and insertions in the 16S rRNA of C. reinhardii

Position	location	E. coli	tobacco	maize
Deletions		in fig. 4		
79-80	A	12	-	-
87-88	A	4	-	-
218-219	A	8	-	-
363-364	H11a	2	-	2
429-430	B	23	-	-
795-796	D	10	9	10
873-874	H22	2	2	2
1078-1085	E	8	8	8
1215-1218	F	4	4	4
1413-1414	G	3	2	3
Insertions				
1229-1233	F	4	4	4
1399-1401	G	3	3	3

Comparisons are with E. coli (16), tobacco (8) and maize (7). The positions and locations of deletions and insertions of more than one base refer to fig. 3 and 4, respectively. H11a and H22 are helices 11a and 22. The number of deleted or inserted bases is indicated in each case.

As shown in fig. 4, the molecule can be folded into four domains with characteristic helices and loops (labelled as in ref. 4). The differences are as follows (cf. fig. 4) : Helix 2 contains one additional base-pair. Helix 11b is missing because

Figure 3.
Sequence of the non coding strand of the chloroplast 16S rRNA gene region of C. reinhardii. The 16S rRNA gene starts and ends at positions 1 and 1493, respectively. The C. reinhardii (upper row) and tobacco (lower row, 8) gene sequences are compared. Identical bases are indicated by dots. Deleted bases are marked with X. The numbered positions do not coincide with the bases of the 16S rRNA since both insertions and deletions have been taken into account. Two additional tobacco insertions are indicated by Δ at positions 795 (9 bases) and 873 (2 bases). The complementary sequences near both ends of the gene are underlined.

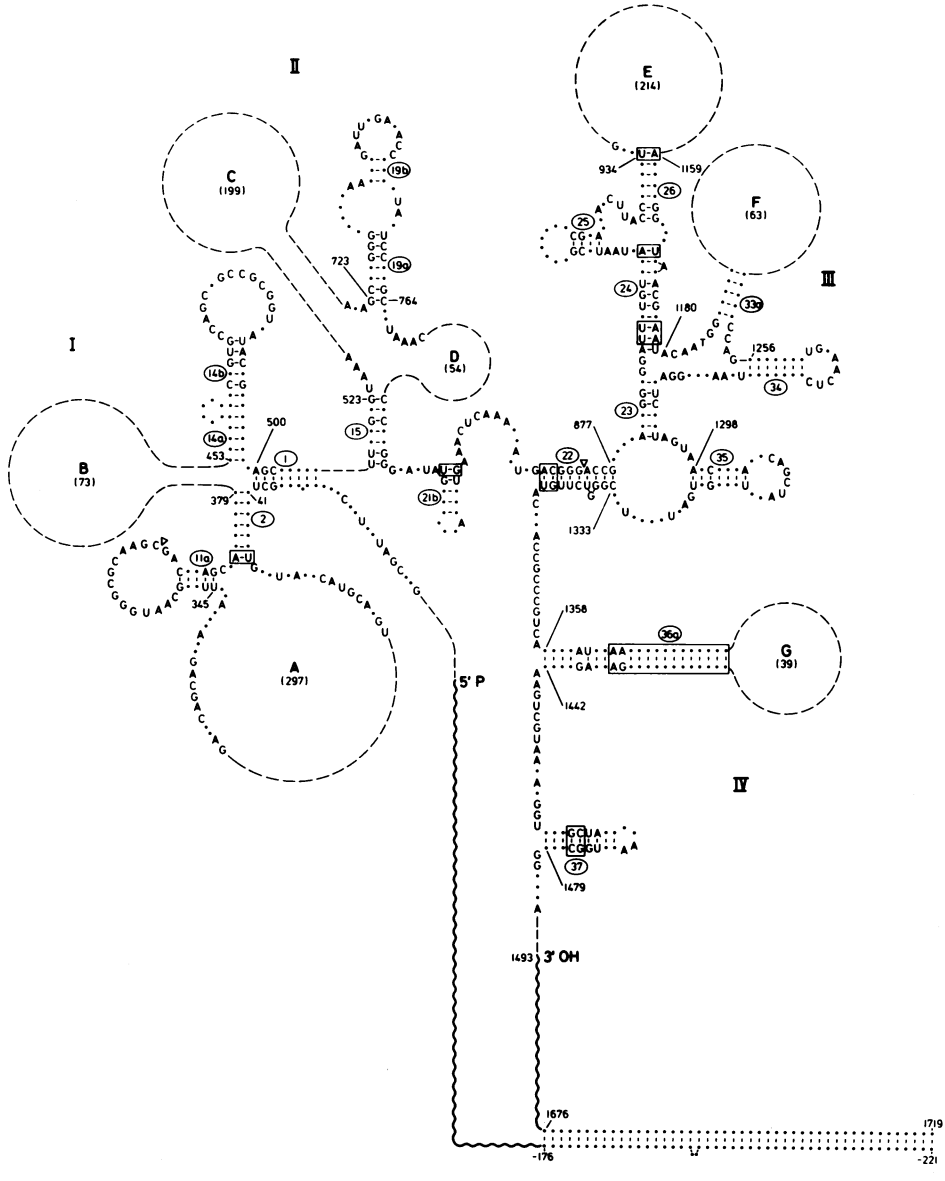

Figure 4.
Secondary structure of the chloroplast 16S rRNA of C.reinhardii. The sequence in fig. 3 has been folded according to the model proposed by Stiegler et al. (fig. 6 of ref. 4). The same nomenclature for the helices (circled numbers) and variable domains has been used. Only the major secondary structure features are shown. Nucleotides which have been numbered as in fig. 3, are indicated by dots and

base-pairs by bars. I, II, III and IV represent the four structural domains. The number of nucleotides in the variable domains A - G is indicated in parenthesis. Nucleotides which are identical or compatible with those of the general model are indicated. Modifications relative to the general model are framed. Deletions in the conserved regions are indicated by ∇. Two regions of the 16S rRNA are drawn in more detail relative to the general model. First, the sequence between bases 349 and 366 shows clearly that helix 11b of the general model does not exist in C. reinhardii. Second, bases 1183 to 1186 leading into helix 33a and bases 1252 to 1255 leading into helix 34 are shown in order to demonstrate the structural resemblance of this region with the binding site for S7 in E.coli 16S rRNA (cf. text). Wavy lines refer to the AT rich segments between the 5' and 3' ends of the 16S rRNA gene and the complementary surrounding sequences.

of a small 2 base deletion (Table I) and is replaced by a larger loop. This is also true for tobacco (8) but not for maize (7). Helices 21b and 22 contain 1 and 2 additional bp, respectively. The structure of helix 24 is slightly different and helix 26 ends with a U-A rather than with a C-G base-pair. Helix 36a contains 12 additional bp with a smaller G domain and the structure of helix 37 is slightly different.

Several deletions occur in the chloroplast 16S rRNA of C. reinhardii as compared to E. coli 16S rRNA. These deletions are also found in the 16S rRNA of maize (7) and tobacco (8) and they occur in the first portion of the RNA (fig. 4, domains A and B, Table I). Other deletions are unique to C. reinhardii 16S rRNA and they are distributed over the second half of the RNA (fig. 4, domains D, E, F and G, Table I). Two small insertions are located in the domains F and G (fig. 4, Table I). With exception of the two small deletions located close to helix 11a and in helix H22, all deletions and insertions occur in the variable domains of the 16S rRNA.

Recently Nomura et al (23) have examined the sequence and secondary structure of the binding sites of the ribosomal proteins S4 and S7 in the 16SrRNA of E. coli. These binding sites are structurally related to the regions of the mRNAs of the ribosomal protein operons where these proteins act as translational repressors (23). It is noteworthy that the S4 binding site which includes helices 1, 13a, 14b and the

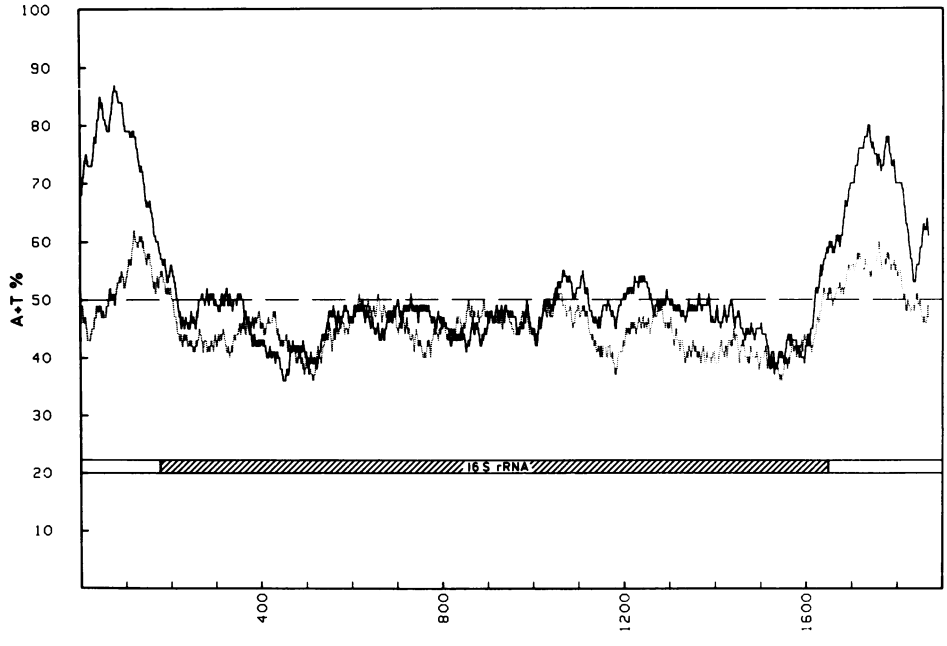

Figure 5.
AT content of the chloroplast 16S rRNA gene regions of C. reinhardii(-) and tobacco (----). Each point plotted on the graph represents the A+T percent in a fragment of 50 bases centered around this point. The lower part of the figure shows the limits of the 16S rRNA gene. The scale is indicated in bp.

adjacent loop and single-stranded region (fig. 4), has been highly conserved between E. coli (23), tobacco (8) and C. reinhardii. Similarly, the binding site of S7 which corresponds to helices 24, 33a, 34 and the surrounding single stranded region (fig. 4) has also been conserved to some extent between E. coli (22) and C. reinhardii. These observations suggest that the sequence homology detected between prokaryotic and chloroplast rRNAs may also extend to some ribosomal proteins. Indeed it has been shown that chloroplast and E. coli ribosomal subunits share several antigenic determinants (24). Recently the gene of the chloroplast elongation factor Tu has been identified in C. reinhardii based on its structural relatedness to its E. coli counterpart (25).

ACKNOWLEDGEMENTS

We thank J. Erickson for helpful comments, F. Veuthey for help with the computer analysis and O. Jenni for drawings and photography. This work was supported by grant 3.659.0.80 from the Swiss National Science Foundation. M.D. was supported by a long term EMBO fellowship.

+To whom correspondence should be addressed

*Present address: Université de Paris-Sud, Centre d'Orsay Biologie Moléculaire Végétale, 91405 Orsay, France

REFERENCES

1. Noller, H.F. and Woese, C.R. (1981) Science 212, 403-411.
2. Stiegler, P., Carbon, P., Zuker, M., Ebel, J.P. and Ehresmann, C. (1981) Nucleic Acid Research 9, 2153-2172.
3. Zweib, C., Glotz, C. and Brimacombe, R. (1981) Nucleic Acid Research 9, 3621-3640.
4. Stiegler, P., Carbon, P., Ebel, J.-P. and Ehresmann, C. (1981) Eur. J. Biochem 120, 487-495.
5. Woese, C.R., Fox, G.E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B.J. and Stahl, D. (1975) Nature 254, 83-85.
6. Woese, C.R., Magrum, L., Gupta, R., Siegel, R.B., Stahl, D.A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J. and Noller, H.F. (1980) Nucleic Acid Research 8, 2275-2293.
7. Schwarz, Zs. and Kössel, H. (1980) Nature 283, 739-742.
8. Tohdoh, N. and Sugiura, M. (1982) Gene 17, 213-218.
9. Rochaix, J.-D. and Darlix, J.-L. (1982) J. Mol. Biol. 159, 383-395.
10. Rochaix, J.-D. and Malnöe, P.M. (1978) Cell.
11. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boxer, H.W., Crosa, J.H. and Falkow, S. (1977) Gene 2, 95-113.
12. Katz, L., Kingsbury, D.T. and Helinski, D.R. (1973). J. Bacteriol. 114, 577-591.
13. Smith, H. and Birnstiel, M.L. (1976) Nucleic Acids Research 9, 2387-2398.
14. Maxam, A. and Gilbert, W. (1980) In Methods of Enzymol. 65, 499-560, Grossman, L. and Moldave, K. eds. Academic Press, N.Y.
15. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
16. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Nat. Acad. Sci. USA 75, 4801-4805.
17. Schwarz, Z., Kössel, H., Schwarz, E. and Bogorad, L. (1981) Proc. Nat. Acad. Sci. USA 78, 4748-4752.
18. Young, R.A. and Steitz, J.A. (1978) Proc. Nat. Acad. Sci. USA 75, 3593-3597.

19. Orozco, E.M., Rushlow, K.E., Dodd, J.R. and Hallick, R.B. (1980) J. Biol. Chem. 255, 10997-11003.
20. Shine, J. and Dalgarno, L. (1974) Proc. Nat. Acad. Sci. USA 71, 1342-1346.
21. Dron, M., Rahire, M. and Rochaix, J.-D. (1982) J. Mol. Biol., in press.
22. Sollner-Webb, B. and Reeder, R.H. (1979) Cell 18, 485-499.
23. Nomura, M., Yates, J.L., Dean, D. and Post, L.E. (1980) Proc. Nat. Acad. Sci. USA 77, 7084-7088.
24. Schneeman, R. and Surzycki, S.J. (1979) Molec. Gen. Genet. 176, 95-104.
25. Watson, J. and Surzycki, S.J. (1982) Proc. Nat. Acad. Sci. USA 79, 2264-2267.