The cloned rRNA genes of *P. lophurae*: a novel rDNA structure

Thomas R. Unnasch and Dyann F. Wirth

Department of Tropical Public Health, Harvard School of Public Health, Boston, MA 02115, USA

Received 9 September 1983; Accepted 27 October 1983

ABSTRACT

The detailed structure of two ribosomal DNA (rDNA) clones CL-1 and HA-2, from the avian malaria parasite *Plasmodium lophurae* has been examined using hybridization and electron microscopy. The results demonstrate that the clone CL-1 contains two regions homologous to 25s rRNA of approximately 2200 base pairs (bp) and 450 bp in length, separated by a non homologous region of 240 bp. CL-1 also contains two regions of approximately 1100 bp and 550 bp homologous to 17s rRNA, separated by a non homologous region of 230 bp. The clone HA-2 contains a single region of 670 bp, which is homologous to 25s rRNA. This region is flanked by non homologous stretches of DNA 940 bp and 110 bp in length. As HA-2 is known to be adjacent to CL-1 in the genome (1), these results suggest that the 25s rDNA is interrupted twice, and the 17s rDNA once, by stretches of DNA not found in mature rRNA.

INTRODUCTION

As a beginning to the study of gene expression during two distinct sections of the life cycle of the malaria parasite, we have been examining the structure of the ribosomal RNA (rRNA) genes in the avian malaria parasite, *Plasmodium lophurae*. We have previously reported the cloning of two classes of the rRNA genes of *P. lophurae*, and on the study of the remaining two distinct, as yet uncloned classes (1). In that work, we have demonstrated that the rRNA genes of this avian parasite are unique in at least two respects. Like the mammalian malaria *P. berghei* (2,3), the genome of this parasite seems to contain very few (7-9) rRNA genes. In addition, these genes are not organized into an easily recognizable tandem array. The work reported here demonstrates that the cloned rDNA classes are unique in at least one additional respect. Both the large (25s) and small (17s) rRNA coding regions are interrupted. The coding region of the 25s rDNA is interrupted twice by DNA not homologous to mature 25s rRNA. The 17s rDNA is also interrupted, with one stretch of DNA that is not homologous to mature 17s rRNA.

MATERIALS AND METHODS

Parasites, Plasmids, and RNA

Saponin freed parasites and purified *P. lophurae* ribosomes were a gift of Dr. Irwin Sherman, University of California, Riverside. The construction and preparation of the plasmids and the isolation of purified DNA fragments has been previously described (1).

Preparation and Hybridization of Southern Blots

Southern blots were prepared from agarose minigels as has been previously described (1), with the exception that the acid wash step was omitted. The preparation of radioactively labelled *P. lophurae* rRNA has also been described (1).

R-Looping and Electron Microscopy

The plasmid CL-1, digested with either Bam HI or Pvu II (see Results), was prepared for R-looping by crosslinking the DNA with 4,5',8 trimethylpsoralen as described by Kaback, et al (4). R-loop reactions consisted of 100 mM pipes (pH 7.2), 500 mM NaCl, 10 mM EDTA, 70% (V/V) formamide, 200 ng *P. lophurae* RNA and 200 ng crosslinked CL-1 DNA, in a total volume of 50 μ l. The R-loop reactions were allowed to proceed for 3 hours at 53°C, and the samples were then quick chilled on ice. Deionized glyoxal was added to a final concentration of 1 M and the samples incubated at 12°C for two hours. The samples were spread using the aqueous technique described by Kolodner (5). The resulting grids were shadowed in an Edwards Rotary Vacuum Coater, and examined at 10,000 x magnification on a Zeiss EM 10 Electron Microscope. Molecule measurements were made from the enlarged negatives using a Summagraphics digitizer and the ADZ custom software micrograph measuring program in conjunction with an Apple II computer.

RESULTS

Restriction Mapping of the Coding Regions of CL-1 and HA-2

The detailed restriction map (Figure 2) previously constructed from studies of the rDNA clones CL-1, HA-1, HA-2 and HB-1 (1) has made it possible to accurately identify the portions of the clone CL-1 homologous to mature rRNA. In order to accomplish this, the DNA insert from either the recombinant plasmid CL-1 or from independently isolated subclones of portions of CL-1 (HA-1 and HB-1) was digested with a series of restriction enzymes, and the fragments separated by agarose gel electrophoresis. The DNA was then blotted onto nitrocellulose, and the filter probed with total 32 P labelled rRNA. The data from these types of experiments allowed the boundaries of the coding regions

Figure 1. Determination of the boundaries of the rRNA homologous regions of CL-1 and HA-2.

The left half of the figure shows the ethidium bromide staining pattern of the agarose gel, while the right half shows the hybridization pattern of the Southern Blot prepared from the same gel probed with radioactive rRNA. Lane A1 is a Hae III digest of pure HA-1 fragment and Lane A2 is a Hinf I digest of the same fragment. Lane B1 is a Hinf I digest of pure CL-1b fragment. Lane C1 is an Msp I digest of pure HB-1 fragment. Lane C2 is an Hinf I digest of the same DNA, and Lane C3 is a Hinf I digest of total HB-1 plasmid. Lane D1 is a Sau 3A digest of pure HA-2 fragment, and Lane D2 is a Sau 3A digest of total HA-2 plasmid. (Although in the hybridization data shown here it is not visible, longer exposures demonstrate that the smallest band in Lane A2 does hybridize to rRNA, while the 350 bp band does not, with a long overexposure. In a similar manner, longer exposures show that the two smallest bands in Lane C2 do hybridize to rRNA, while the 250 bp band does not. The low intensity of hybridization of these small bands is probably due to the small size of the rRNA homologous region they necessarily contain, and the poor transfer efficiency of small fragments to nitrocellulose.)

found in the cloned rDNA unit represented by HA-2, C1-1, and the independently isolated subclones of CL-1 to be delineated quite accurately. For example, the data in Lane A1 and Lane A2 of Figure 1 demonstrate that the left hand boundary of the 17s RNA coding region found in the subclone HA-1 falls between the leftmost Hinf I site and the leftmost Hae III site of the cloned DNA. As Lane A2 demonstrates, the terminal 350 bp Hinf I fragment (the second largest) does not hybridize to rRNA. The data in Lane A1 shows that the 400 bp Hae III terminal fragment (the second largest) does hybridize weakly to rRNA. This localizes the leftmost boundary of the coding region to the 50 bp between these sites.

In a similar manner, the data in Lanes A1 and A2 localize the right hand boundary of the 17s coding region of HA-1 to the 100 bp between the right-most Hinf I site and the right-most Hae III site of this cloned DNA. The 230 bp

Figure 2. A summary of the regions of HA-2 and CL-1 homologous to mature rRNA.

The hatched bars under the restriction map indicate those regions which are homologous to mature rRNA. Each bar is also labelled with the type of rRNA with which it is homologous (1). The scale is in bp, with the left most CLA-1 site of CL-1 labelled as 0. (No sites for BamH I, Pvu II, or EcoR 1 are present in any of the clones.)

terminal Hae III fragment (the smallest in lane A1) does not hybridize to rRNA, but the 330 bp terminal Hinf I fragment (the third largest band in Lane A2) hybridizes weakly. The hybridization data therefore demonstrate that the rDNA contained in HA-1 consists of an 1120 bp region homologous to 17s rRNA, flanked on both sides by regions that are not homologous to any mature rRNA.

In a similar manner, the data shown in Lane B1 demonstrates that the central 1.8 kb Hind III fragment of CL-1 (CL-1b) contains two distinct coding regions separated by 840 bp of non-coding DNA. As is shown in Figure 1, Lane B1, one of the five fragments (the second largest) produced by a Hinf I digest of CL-1b does not hybridize to rRNA. This fragment is positioned between the 220 bp Hinf I doublet, which hybridizes to 25s rRNA, and the 640 bp and 230 bp Hinf I fragments, which hybridize to 17s rRNA (1). The second largest Hinf I fragment thus represents part of the spacer between the 17s and 25s rRNA genes. Similar mapping experiments with Sau 3A/Hinf I double digests place the right hand boundary of the 17s coding region to within 100 bp of the single Sau 3A

site of CL-1b. The 25s and 17s rRNA coding regions of CL-1b are thus separated by a spacer approximately 840 bp in length.

Lanes C1, C2, and C3 delineate the boundaries of the 25s coding region in HB-1. The data in lane C-1 demonstrates that the 360 bp Msp 1 fragment of HB-1 hybridizes weakly to rRNA, suggesting that only part of the fragment codes for rRNA. In fact, Lane C2 demonstrates the 250 bp Hinf 1 terminal fragment (the 5th largest), which is contained within the 360 bp Msp 1 fragment does not hybridize to rRNA. Additional confirmation that this 250 bp Hinf 1 fragment does not hybridize to rRNA is shown in Lane C3. In this case, the 250 bp Hinf 1 terminal fragment is linked to a portion of pBR322 to produce a chimeric fragment of approximately 900 bp. This fragment, indicated by an arrow, still does not hybridize to rRNA. In summary, the clone HB-1 contains a stretch of approximately 2100 bp of DNA homologous to rRNA. This region is separated from the region of the central 1.8 kb Hind III fragment homologous to 25s rRNA by approximately 240 bp of non-homologous DNA.

Finally, the region of clone HA-2 homologous to a mature rRNA has been determined. Although this clone is not part of CL-1, but as previously shown, it is adjacent to the 25s coding end of CL-1 within the genome of P. lophurae (1). It also codes for 25s rRNA (1) and thus probably represents the 25s terminus of the rDNA unit. Lanes D1 and D2 contain the information needed to delineate the boundaries of the coding region of this clone. As the data in lane D-1 demonstrates, only two of the four fragments produced in a Sau 3A digest of HA-2 insert hybridize to rRNA. The largest (930 bp) fragment, and the smallest (110 bp) band do not hybridize. These two fragments represent the terminal Sau 3A fragments of HA-2. It therefore appears that the cloned insert HA-2 contains a stretch of about 670 bp homologous to rRNA. This homologous region is flanked on both sides by non-homologous DNA, of 930 bp and 110 bp in size. To further confirm that the 110 bp fragment does not hybridize to rRNA, a chimeric fragment of 350 bp was generated by a Sau 3A digest of the complete HA-2 plasmid. This fragment, indicated by the arrow in Lane D2, contains 240 bp of pBR322 DNA. As can be seen, this fragment also does not hybridize to rRNA.

A summary of the data discussed above is presented in Figure 2. The restriction data suggests that the coding regions for both the 17s and 25s rRNA found in CL-1 and HA-2 are interrupted by DNA that is not homologous to mature rRNA. The 17s gene is split into 1120 bp and 540 bp segments by a non-homologous region of 230 bp. Likewise, the 25s gene is divided; once within CL-1 into segments of 2100 bp and 450 bp by a 240 bp interruption, and

Figure 3. R-loop analysis of the 17s rDNA of CL-1.

R-Loops were formed with CL-1 digested with Bam HI, as described in Materials and Methods. The arrow highlights the stem structure discussed in the text.

once at the CL-1 - HA-2 boundary by a 110 bp interruption. This last interruption produces the terminal 670 bp coding segment found in HA-2.

R-Loop Analysis of CL-1

Further evidence demonstrating the presence of interruptions within the 17s and 25s rRNA genes of the plasmid CL-1 was obtained from R-loop analysis. In the first such experiment, the plasmid was digested to completion with BamH I. This enzyme does not cut within the CL-1 insert (1), and cuts the vector pBR322 DNA only once. This results in a linear molecular containing 352 bp of pBR322 at the 25s end of the insert, and the remaining 4010 bp of pBR322 connected to the 17s end of the insert. This DNA was then incubated with RNA under R-loop conditions, and spread for electron microscopy, as described in materials and methods. An example of the class of molecules containing R-loops is shown in Figure 3. The R-loop formed by the 17s rRNA is clearly visible. Within this R-loop, a small stem is present, projecting from the double stranded side of the loop. (In this particular molecule, as in many others seen, the 25s rRNA R-loop is present as an open fork, with the double stranded side of the R-looped DNA also ending in a fork. This second fork probably consists of the unhybridized pBR322 tail, and that portion of the 25s rRNA

Figure 4. A summary of the contour lengths of the 17s rDNA R-loops.

Each measurement is expressed as a mean value in bp (for double stranded regions) or bases (for single stranded regions), together with the sample standard deviation. The number in parenthesis is the total number of measurements in each class. The column labelled "predicted" contains the contour lengths predicted by the hybridization data.

homologous to HA-2.) Contour measurements of the R-loop containing molecules are presented in Figure 4. These results demonstrate that the R-loop observed corresponds in both size and position to the size and position of the 17s rDNA, as predicted by the hybridization data. In addition, the stem seen is in a position which allows it to be assigned to the 230 bp interruption in the 17s rRNA gene detected by the hybridization data.

In order to visualize the interruption in the 25s rDNA of CL-1, the experiment was repeated using CL-1 plasmid digested with Pvu II. This enzyme does not cut within the insert, and cuts the vector DNA once. This results in a linear molecule having 2042 bp of pBR322 DNA attached to the 25s end of the insert, and 2320 bp of pBR322 DNA attached to the 17s end. This DNA was then used to produce R-loops as described above, and the resulting mixture was spread for electron microscopy. An example of the class of molecules containing R-loops is shown in Figure 5. Both the 17s and 25s R-loops are clearly visible, although the 17s R-loop is broken. In addition, a stem on the double stranded side of the 25s R-loop is seen, much like that seen on the 17s R-loop presented above. As predicted, a small tail of rRNA is seen to extend from the fork of the 25s R-loop distal to the 17s R-loop. This probably corresponds to the rRNA homologous to the insert HA-2. Measurements of the sections of several such molecules are presented in Figure 6. This data demonstrates that both the 25s and 17s R-loops correspond both in position and size to the predictions provided by the hybridization data. The stem seen in

Figure 5. R-loop analysis of the 25s rDNA of CL-1.
R-Loops were formed with CL-1 plasmid digested with Pvu II, as described in Materials and Methods. The arrow highlights the stem structure discussed in the text.

Figure 6. A summary of the contour lengths of the 25s rRNA R-loops.
Each measurement is expressed as a mean value in bp (for double stranded regions) or bases (for single stranded regions) together with the sample standard deviation. The number given in parenthesis is the total number of measurements taken in each class. The column labelled "predicted" contains the contour lengths predicted by the hybridization data.

Figure 7. A summary of the position of the coding regions in the cloned rDNA unit of *P. lophurae*.

the 25s loop is also found to correspond in position to the position of the 240 bp non-homologous region detected by hybridization. The R-loop data, with the hybridization data presented above, demonstrate that the 25s rDNA and the 17s rDNA of CL-1 is interrupted by DNA that is not homologous to mature rRNA. It is likely, therefore, that the complete rDNA unit represented by the clones HA-2 and CL-1 contains three interruptions; two in the 25s rDNA and one in the 17s rDNA. A summary of the structure of the cloned rDNA unit is presented in Figure 7.

DISCUSSION

The results presented in this paper demonstrate that both the 17s and 25s rDNA contained in the clone CL-1 are interrupted by DNA which does not hybridize the mature rRNA. The 17s gene is interrupted once by 230 bp of non-coding DNA, while the 25s gene is interrupted once by 240 bp of non-coding DNA. It is also likely that a second interruption of at least 110 bp is present in the 25s rDNA of the cloned rDNA unit, at the junction of the clones HA-2 and CL-1. Although the existence of this interruption is supported only by hybridization data, this data alone strongly suggests that this interruption exists. This is because the 350 bp pBR322-insert chimeric fragment does not hybridize to total rRNA (Figure 1, Lane D2). In contrast, the 400 bp Hpa II fragment of HA-1, which, as discussed above, can contain at most 50 bp of coding DNA, hybridizes quite clearly. This means that there cannot be more than 50 bp of homologous DNA in the 110 bp Sau 3A end terminal fragment of HA-2. Therefore, some interruption in the coding sequence must exist.

In some organisms, particularly the *Diptera*, two interruptions are found within the large rRNA coding region (6,7,8). One of these interruptions results from the nicking and removal of 140 bases from the large rRNA during processing. Although such nicking does occur in some species of *Plasmodia* (2,9), this cannot explain either of the two interruptions found in *P.*

lophurae. First, denaturing gel electrophoresis of total parasite RNA shows no evidence for such nicking (1). This result makes P. lophurae similar to mammalian malarial P. falciparum (10) and P. knowlesi (2) where such a nick does not appear to exist. Even if a nick existed near the end of the large rRNA that was not detected by the denaturing gel, it could not explain the two interruptions described here. This is because purified 25s rRNA from denaturing gels hybridizes to all three coding regions in question (1). If one of the non-coding regions was generated by post-translational processing, one would not have expected the smaller coding region distal to the cut to hybridize to purified, high molecular weight RNA. Finally, the R-loop data presented above also suggest that the 25s rRNA is contiguous across the 240 bp interruption found in CL-1, since a nick in the mature 25s rRNA in this position would not be expected to result in the stem structure seen in these experiments.

A unique characteristic of the rDNA of P. lophurae, as represented by the cloned fragment contained within CL-1, is the presence of a non-coding interruption in the 17s rDNA. Since the 5.8s rRNA gene of most organisms examined to date is found separating the large and small rRNA genes (8,11,12,13) one must question if the smaller region assigned to the 17s rDNA, is in fact homologous to 5.8s rDNA. This is unlikely for several reasons. First, this region has been shown to hybridize to purified 17s rRNA (1). If the region was in fact homologous only to 5.8s rRNA, this hybridization could only be explained by massive contamination of the purified 17s rRNA with 5.8s rRNA. Such contamination is not seen when the purified 17s rRNA is run on a denaturing gel (data not shown). Secondly, such an explanation cannot account for the stem structure seen in the R-loop experiments. Finally, the length of 5.8s rRNA is known by sequence analysis to be always near 160 bases in length (4). By both restriction enzyme analysis and R-looping, it is clear that the small coding region of the 17s rDNA is much larger than this. When taken together, these three facts make it very unlikely that the smaller coding region assigned to the 17s rRNA is in fact coding for 5.8s rRNA.

The interruptions seen in CL-1 are not due to small additions and or substitutions during cloning. As previous work has demonstrated (1), CL-1 is one of four independently isolated rDNA clones, all of which have basically identical restriction maps. In addition clones of the two regions of CL-1 containing the interruptions have been obtained by a totally independent cloning methodology directly from genomic DNA (1). These clones have identical restriction maps and hybridization patterns when compared to the corresponding

regions of CL-1. The presence of 6-7 identical, independently obtained isolates of each of the regions containing the interruptions renders any cloning artifact very unlikely.

A final question which may be raised concerning the interruptions in the rDNA of CL-1 and HA-2 concerns the expression of the rRNA genes in the malaria parasite. It is possible that CL-1 and HA-2 may represent rDNA that is not normally expressed during the erythrocytic stages of the parasite life cycle. Thus, the proposed interruptions may represent DNA that is not homologous only to erythrocytic stage rRNA, and is homologous to some other form of the rRNA produced by *P. lophurae* during its life cycle. We feel that this is unlikely, since the R-loop data clearly demonstrates that the interruptions represent additions rather than substitutions of sequence when the rDNA is compared to the mature rRNA. This is because only stem type structures were seen in the R-loop experiments, and not the open single loop structures expected if substituted regions existed between the cloned rDNA unit and the hybridizing rRNA. However, the data derived from southern blots of genomic DNA double cut with CL-1 and Hind III demonstrates that no additional DNA is inserted into any of the known DNA classes of *P. lophurae* (1). The Hind III sites fall near the boundaries of the interruptions, and this experiment would have easily detected the size changes produced by the hypothesized additions. We conclude that the rDNA represented by CL-1 and HA-2 must contain interruptions. The question remains whether the rRNA genes containing interruptions are expressed. In several lower eucaryotes, including *Tetrahymena thermophila* (15) and *Physarum polycephalum* (13) intron containing rRNA genes appear to be transcribed. In contrast, the intron containing rDNA of *Drosophila melanogaster* appears to be transcriptionally inactive (16,17). Although the fact that *P. lophurae* contains few rRNA genes (1) suggests that the classes of rDNA represented by the four CLA1 clones we have isolated will be transcribed, further studies are necessary. The current availability of intron specific subclones will prove helpful in resolving this question.

Finally, it should be noted that the sizes of the coding regions seen in the R-loop experiments are somewhat larger than those predicted by the hybridization data. Although some of this difference may be ascribed to variations in contour length caused by the spreading conditions (e.g. the pBR322 stems are about 5% longer than predicted), most of the difference is probably due to the way in which this map was generated. All restriction fragments generated during the mapping experiments were equalized to the minimum estimate of the size of the cloned fragment from which the map was

derived. As this figure was usually less than the sum of the fragments, some underestimation of fragment sizes probably occurred. This would in turn lead to an underestimation of the sizes of the coding regions.

ACKNOWLEDGEMENTS

We would like to thank Irwin Sherman for providing us with parasites, Richard Kolodner and Paul Morrison for help with the electron microscopy, and New England Biolabs for a generous gift of material. This work was supported by a grant to Dyann F. Wirth from the National Institutes of Health. Dyann F. Wirth is a recipient of a special award in molecular parasitology from the Burroughs Wellcome Foundation. Thomas R. Unnasch was supported under a National Research Service Award #1 F32 AI05667-01A1 from the N.I.A.I.D.

REFERENCES

- 1) Unnasch, T.R., and Wirth, D.F. (1983) Nuc. Acids Res., submitted.
- 2) Dame, J.B., and McCutchan, T.F. (1983) Mol. Biochem. Parasit. 8, 263-280.
- 3) Dame, J.B., and McCutchan, T.F. (1983) J. Biol. Chem. 258, 6984-6991.
- 4) Kaback, D.B., Angerer, L.M., and Davidson, N. (1979) Nuc. Acids Res. 6, 2499-2516.
- 5) Kolodner, R. (1980) Proc. Nat. Acad. Sci. U.S.A. 77, 4847-4851.
- 6) Glover, D.M., and Hogness, D.S. (1977) Cell 10, 167-176.
- 7) Wellauer, P.K., and Dawid, I.B. (1977) Cell 10, 193-212.
- 8) Pellegrini, M., Manning, J., and Davidson, N. (1977) Cell 10, 213-224.
- 9) Miller, F.W., And Ilan, J. (1978) Parasitology 77, 345-365.
- 10) Hyde, J.E., Zolg, J.W., and Scaife, J.G. (1981) Mol. Biochem. Parasit. 4, 283-290.
- 11) Leon, W., Fouts, D.L., and Manning, J. (1978) Nuc. Acids Res. 5, 491-504.
- 12) Boseley, P.G., Tuyns, A., and Birnsteil, M.L. (1978) Nuc. Acids Res. 5, 1121-1137.
- 13) Campbell, G.R., Littau, V.C., Melera, P.W., Allfrey, V.G., and Johnson, E.M. (1979) Nuc. Acids Res. 6, 1433-1447.
- 14) Erdmann, V.A., Huysmans, E., Vandenberghe, A., and DeWachter, R. (1983) Nuc. Acids Res. 11, r105-r133.
- 15) Cech, T.R., and Rio, D.C. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 5051-5055.
- 16) Long, E.O., and Dawid, I.B. (1979) Cell 18, 1185-1196.
- 17) Long, E.O., Collins, M., Kiefer, B.I., and Dawid, I.B. (1981) Mol. Gen. Genet. 182, 337-384.