Volume 9 Number 12 1981 Nucleic Acids Research

Sequence of a yeast DNA fragment containing a chromosomal replicator and a $tRNA^{Glu}_3$ gene

Horst Feldmann, Judith Olah+ and Hildegard Friedenreich

Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, Goethestrasse 33, 8000 München 2, GFR

Received 12 March 1981

ABSTRACT

The sequence of a 1.9 kb Bam x Hind III fragment from yeast has been determined. This fragment is part of a yeast 6.7 kb Hind III segment cloned into pBR322 (pY20). The fragment carries a single gene for a glutamate tRNA which has no intron. According to genetic analyses [1] this fragment also contains a yeast chromosomal replicator. We have analyzed the sequence for potential open reading frames and for several structural features which are thought to be involved in the initiation of DNA replication. Hybridization studies have revealed that portions of this sequence are repeated within the yeast genome.

INTRODUCTION

For a study of the organization and expression of yeast tRNA genes we have isolated a number of plasmids [2] . Previously we reported the structure of a $tRNA^{Met}_3$ gene [2]; we have since analyzed the DNA sequences of several other tRNA genes [3]. Here we wish to report the characterization of a 1.9 kb fragment from clone pY20 which is shown to contain a single $tRNA^{Glu}_3$ gene. More interestingly, genetic analyses by C. Cummins and M. Culbertson [1] revealed that this fragment carries a putative chromosomal replicator. We therefore have analyzed and compared this DNA sequence not only with that for another yeast $tRNA^{Glu}_3$ gene from our clone pY5 [4], but also with other putative yeast replicators [5,6].

MATERIALS AND METHODS

Chemicals and enzymes. [^{32}P] phosphate (carrier-free) was purchased from New England Nuclear Corp. $Na^{125}I$, high specific activity, was obtained from the Radiochemical Centre, Amersham. SeaKem and Seaplaque agarose were products of Marine Colloids,

Rockland. Acrylamide and N,N'-Bismethylenacrylamide were purchased from Serva, Heidelberg. Other chemicals were of analytical grade. Total yeast tRNA was obtained from Boehringer GmbH, Mannheim. Some individual yeast tRNA species were a gift of G. Dirheimer and G. Keith, Strasbourg. T4 DNA ligase was a product of Miles Corporation. T4 polynucleotide kinase was from Boehringer GmbH, Mannheim. Restriction endonuclease Taq YI was from Microbiological Research Establishment, Porton; Eco RI, Sau 3A, Sau 96, and Alu I were gifts of R.E. Streeck; Bam HI and Hae III were gifts of U. Hänggi; Hind III [7] and Hpa II [8] were prepared by standard procedures with slight modifications.

Plasmid DNA. Clone pY20 was from a collection of tRNA gene-carrying yeast DNA/pBR322 plasmids constructed as reported in ref. 2. The cells were grown in LB medium in 15 l fermenters [9], and plasmid DNA was prepared by scaling up the procedure of Birnboim and Doly [10].

Preparation of DNA fragments, gel electrophoreses, and restriction mapping were performed in a similar way to that described in ref. 2.

Labeling procedures and hybridizations. Yeast tRNAs were labeled with ^{125}I and purified by the procedures described in ref. 2. Hybridization of [^{125}I]tRNA to DNA fragments transferred to nitrocellulose filters [11] were performed as described earlier [2]. DNA fragments for DNA-DNA hybridizations were 5'-end labeled by the use of [γ-^{32}P]ATP and polynucleotide kinase [12]. Hybridizations were carried out according to ref. 13.

DNA sequencing followed the protocol of Maxam and Gilbert [14]. 8% and 20% gels of 0.5 or 0.2 mm thickness were employed. Partial cleavage at purine residues was achieved by the reaction described by Gray et al. [15]. DNA fragments for sequencing were prepared as described earlier [2].

RESULTS AND DISCUSSION

1. Characterization of pY20 by restriction enzyme analysis and hybridization to tRNA.

Plasmid DNA was prepared from clone pY20 [2], cut with restriction endonucleases Eco RI, Hind III, Bam HI, and combina-

tions of these enzymes and run out on agarose gels. A similar analysis was carried out with the reisolated yeast Hind III segment from this plasmid. From these data the restriction map shown in Fig. 1A was deduced. Hybridization with [^{125}I]tRNAs was observed only with the 1.9 kb Bam x Hind III fragment. Therefore, a fine restriction map of this fragment was determined for the following enzymes: Hae III, Hpa II, Hha I, Alu I, Sau 96, Sau 3A, and Taq YI (Fig. 1B). Several procedures were used for this mapping: digestions with combinations of the enzymes, secondary digestions of isolated subfragments with appropriate enzymes, and digestions of 5'-end-labeled subfragments. In some positions, two sites for the same enzyme are located very close to each other (e.g. Sau 3A, Taq). These sites were only discovered after having established the complete DNA sequence.

Hybridization of the subfragments with [^{125}I]tRNA (not shown)

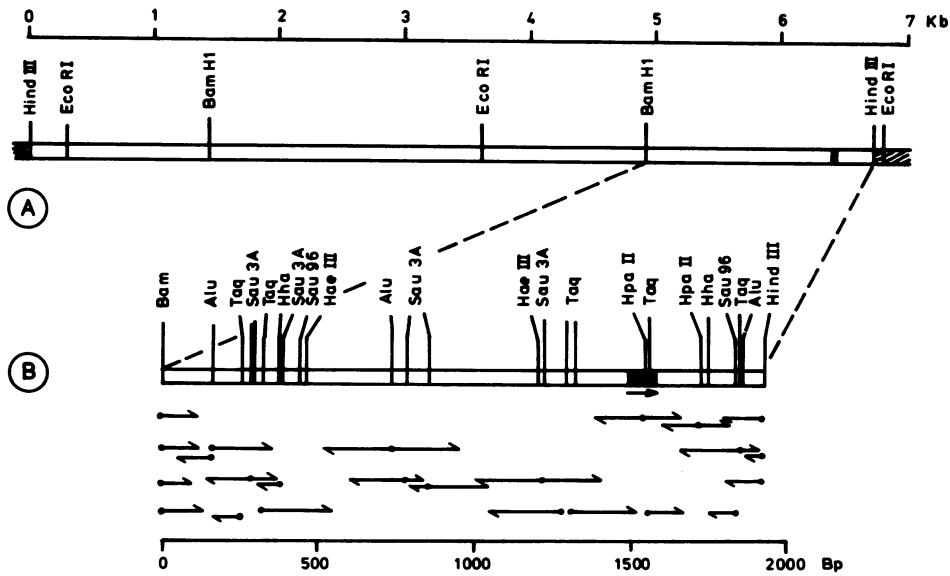

Fig. 1: Restriction endonuclease cleavage map of pY20.
A: The 6.7 kb yeast DNA insert in pBR322 (hatched ends)
B: Fine restriction mapping of the 1.9 kb Bam x Hind III fragment and DNA sequencing strategy. The arrows indicate the direction and extent of sequencing. The position of the tRNA$^{Glu}_{3}$ gene is shown in black; the arrow indicates the direction of transcription.

restricted the location of a tRNA gene to a region near the Hind III site of the 1.9 kb fragment.

2. Sequence of the 1.9 kb Bam x Hind III fragment from pY20.

The sequence of the fragment is presented in Fig. 2. The fragment, bounded by a Bam Hl site and a Hind III site, is oriented as shown. The length of the fragment is 1919 base pairs. The base composition of the upper strand in Fig. 2 is 32.2% A, 17.35% G, 19.44% C, and 31.01% T (the light strand); the (G+C) content of the fragment is 36.8% which is close to that of total yeast DNA (ca. 38%) [16].

The strategy used to determine the DNA sequence is outlined in Fig. 1B. The arrows indicate the extent and direction of sequencing from the various restriction sites. The Maxam and Gilbert protocol [14] was followed; the label and cut method was applied as well as the strand separation of labeled subfragments. In this way, the Hpa II, Alu I, Sau 3A, and Taq fragments were sequenced from both ends (using multiple loadings on 20% and 8% gels, up to 250 bases could be read). This allowed us to deduce most of the sequence from both of the strands. In all cases, however, sequence determinations were performed at least in duplicate.

3. Analysis of the DNA sequence.

A number of computer programs [17] were used to analyze the DNA sequences for restriction sites, base distributions, potential coding regions, particular structural features, and also for comparisons of these features with other DNA sequences. The restriction sites obtained by physical mapping (Fig. 1) were confirmed by sequencing.

(a) The $tRNA^{Glu}_3$ gene

The 1.9 kb fragment contains a single gene for $tRNA^{Glu}_3$ [18], the coding region of which is located between positions 1495 and 1566 (Fig. 2). Transcription is from left to right in this orientation; the tRNA gene has no intervening sequence. A more detailed description of this portion of the DNA sequence and a comparison with another $tRNA^{Glu}_3$ gene in our yeast clone pY5 is presented in the accompanying paper [4]. No other gene for a known tRNA [19] was found in the 1.9 kb fragment by a computer

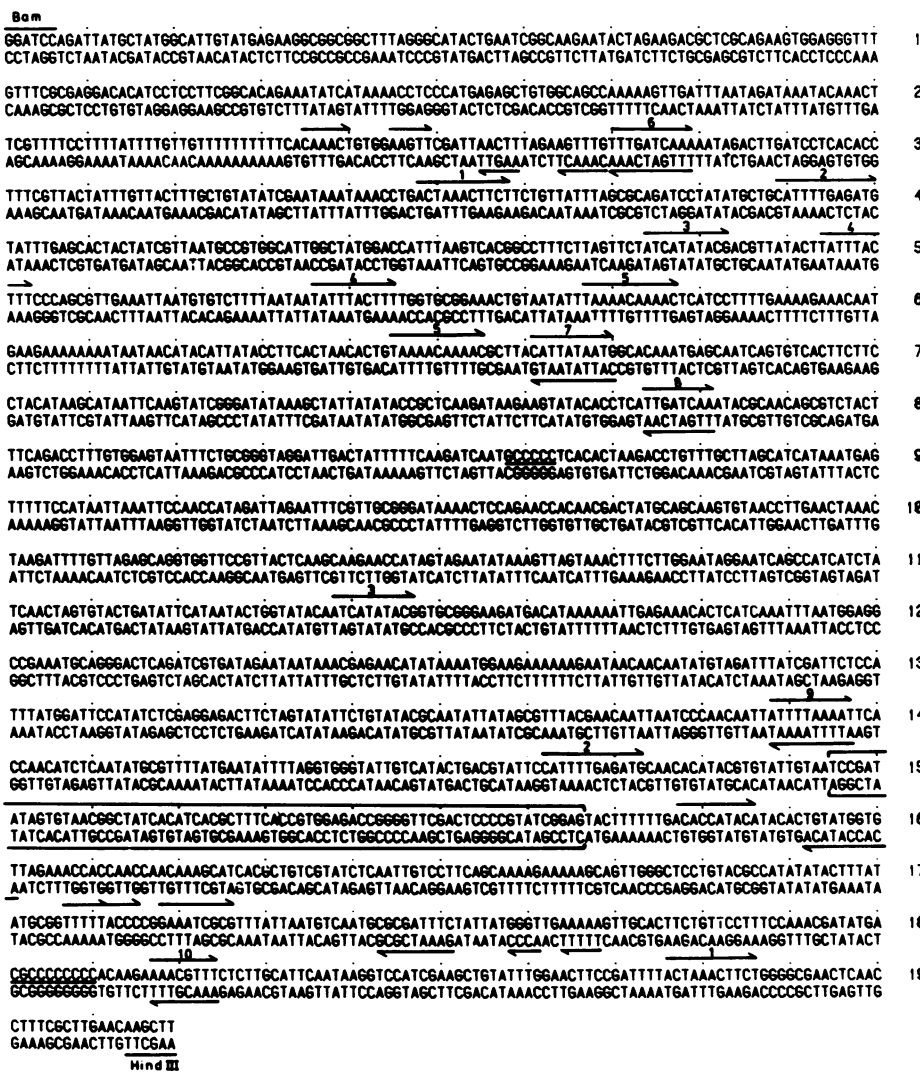

Fig. 2: DNA sequence of the 1.9 kb Bam x Hind III fragment from pY20.
The sequence is oriented in a way so that the transcription of the $tRNA^{Glu}_{3}$ gene (boxed) is from left to right. Numbered arrows above the sequence indicate internal homologies and symmetries. Inverted repeats are pointed out by arrows above and below the sequence (cf. also Fig. 3).

search. Also, no other tRNA gene-like structure could be detected by inspection of the sequence.

(b) Homologies and symmetries

The DNA sequence contains 5 repetitions of ≥10 bases, but no direct repeats of ≥ 10 bases. True palindromes (≥ 8 base pairs) were detected in the following positions (lengths in parantheses): 270-279(10); 661-670(10); 773-780(8); 1388-1397(10); 1817-1824(8). In addition, several inverted repeats were seen, from which possible secondary structures could be formed (see Fig. 2).

(c) Variation of (G+C) content and possible coding regions

The variation of the (G+C) content was computer calculated for every position by averaging the (G+C) content of 20 base pairs each to the left and to the right of this position [17]. In a graphic representation (not shown), the tRNA gene is clearly seen as a peak (> 50% G+C) around position 1550. Other relatively (G+C) rich regions occur around position 1820 and within the first 160 base pairs. The peak around position 1820 is due to a block of 10 (G:C) pairs (see next section).

An analysis to locate all potential initiation and termination codons in the DNA sequences revealed that there are two relatively short open reading frames, one starting at position 1701 (ATG) extending for only 59 amino acids. We do not know, whether mRNA transcription could be initiated so close to a tRNA gene. There is a third reading frame entering the DNA sequence from the beginning that has two adjacent stop codons beyond position 181 (≥ 56 amino acids). This region coincides with a region relatively high in (G+C) content, and it is followed by a very (A+T) rich segment and the possible secondary structure indicated in Fig. 2. The opposite strand has only one open reading frame (1518 to 1832) equivalent to 105 amino acids, but it extends into the tRNA region and therefore is not very likely to code for a protein.

(d) Location of the chromosomal replicator

Genetic analyses carried out by C. Cummins and M. Culbertson have located the 1.9 kb fragment on yeast chromosome V. They have also demonstrated that the fragment contains a putative chromosomal replicator. These experiments will be reported else-

where [1]. From earlier investigations of cloned DNA containing chromosomal replicators [20,21] it seems rather unlikely that they share extended serial homologies, because cross-hybridizations were not observed in these cases. Nevertheless, we checked our sequence, the putative replicator region coupled to the TRP1 gene [5], and the putative origin of replication in the yeast 2μ DNA [6], which is thought to be replicated together with chromosomal DNA [22], for common sequences or other obvious structural similarities. This search was negative for even short ($\geq$ 8 base pairs) homologies in all three sequences.

One characteristic often found in origins of replication is a block of (G:C) pairs [23-25]. A block of 5 (G:C) pairs is seen around position 950 in the TRP1 sequence [5], and two such blocks, one between positions 858 and 864 and the other one between 1801 and 1810 occur in our sequence. A similar block is not found near the putative replicator region in the 2μ DNA; there are only several alternating (G:C) and (C:G) pairs around position 4020 [6]. In the latter case, two alternate potential secondary structures near the putative origin of replication have been suggested [6]. Elements of symmetry which can be drawn as secondary structures are also visible in the vicinity of the (G:C) blocks in the TRP1 sequence [5] (between positions 987 and 1064) and, although less extended and differently shaped, in our sequence (Fig. 3). However, the secondary structures found near the origins of replication of other genomes [e.g. 23-25] are also variant. The secondary structure in Fig. 3 would have a theoretical thermodynamic stability of the order of -28 Kcal $mol\ell^{-1}$ [26], as compared to a value of -50 Kcal $mol\ell^{-1}$ in the case of the yeast 2μ DNA [6].

(e) Genomic redundancy of DNA sequences from pY20

In order to obtain an estimate of the redundancy of $tRNA^{Glu}_{3}$ genes in the yeast genome, we hybridized corresponding ^{32}P-labeled DNA segments to total yeast DNA which had been restricted with Hind III or Eco RI and separated on 0.8% agarose gels. Fig. 4 shows the autoradiographs of this experiment. Whereas several bands appear to comigrate in the Hind III pattern, 14 bands of different lengths are resolved in the Eco RI pattern. This number is in good agreement with the estimates for the re-

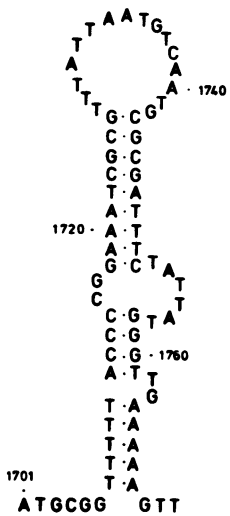

Fig. 3: Possible secondary structure in the DNA sequence in the vicinity of the (G:C) block (pos. 1801-1810).
For convenience, only the folding of the upper strand is shown. The respective area is also indicated by arrows in Fig. 2.

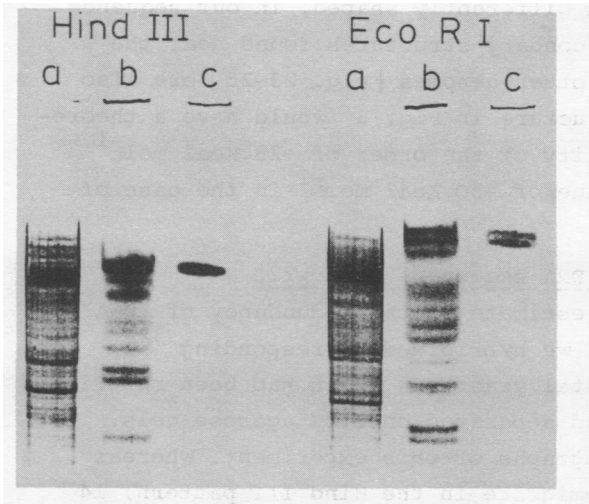

Fig. 4: Hybridization of labeled DNA subfragments from the 1.9 kb Bam x Hind III fragment to total yeast DNA (cf. text) a: restriction patterns (Hind III or Eco RI, respectively) resolved on 0.8% agarose gels; b: radioautogram after hybridizing the Taq x Hpa II fragment (1290-1541); c: radioautogram after hybridizing the Bam x Alu fragment (1-157); virtually the same results were obtained with the following fragments: 157-249; 1-289; 289-375; 375-776; 776-851.

dundancy of other yeast tRNA genes [16].

Since in the case of the $tRNA_3^{Met}$ gene [2] we had observed that sequences outside the coding region gave a number of hybridizing bands, we also performed hybridizations with the labeled subfragments of the 1.9 kb fragment described in the legend to Fig. 4. The example shown in Fig. 4 is representative of all our results. Two hybridizing bands are observed in the Eco RI pattern. This result indicates that sequences in the vicinity of tRNA genes or replicators may be reiterated in the genome. Examples for both, tRNA genes in human DNA [27], and replicators in yeast [28], have been discussed recently. A similar observation that sequences in the vicinity of a tRNA gene may be repeated in the genome has been made in the case of pY5 [4].

CONCLUDING REMARKS

We have analyzed the sequence of a yeast DNA fragment containing a tRNA gene in addition to a putative chromosomal replicator. The results provide further information which might help to define structural requirements for replicating functions. In terms of obvious structural features, we have only discovered the ones presented above. It should be mentioned, that the possible secondary structure found adjacent to the 10 base pair (G:C) block (positions 1801-1810) appears to be significant statistically and shares characteristics with several other putative replicators. No such structure is detectable in the vicinity of the 5 base pair (G:C) block (around position 950) in this same fragment. It is still necessary to delimit the "minimal" region essential as an origin of replication. This can be accomplished by subcloning appropriate fragments and performing further genetic analyses.

We are also continuing work to identify and analyze the putative gene suggested by the open reading frame at the "left" end (Bam H1 site) of the 1.9 kb fragment. The identification of a known structural gene would help to correlate the genetic and physical map.

ACKNOWLEDGEMENTS

We wish to thank Ms. Christa Bleifuß and Ms. Petra Müller

for excellent technical assistance. We are grateful for the gift of enzymes to Drs. R.E. Streeck, U. Hänggi, W. Hörz, and F. Fittler from this laboratory. P. Neumaier from this laboratory has kindly allowed us to use his computer programmes. We thank C. Cummins and M. Culbertson (MIT, Cambridge) for having communicated their results prior to publication. The Deutsche Forschungsgemeinschaft (Forschergruppe Genomorganisation) and the Fonds der Chemischen Industrie have supported this work .

REFERENCES

+ Judith Olah's present address: Dept. Genet. and Cell Biol. Univ. Minn.,250 Biol.Sci.Center, St. Paul, Minn. 55108, USA
1 Cummins, C. and Culbertson, M., manuscript in preparation
2 Olah, J. and Feldmann, H. (1980) Nucl. Acids Res. 8, 1975-1986
3 Feldmann, H. et al., manuscript in preparation
4 Eigel, A., Olah, J., and Feldmann, H. (1981) Nucl. Acids Res.
5 Tschumper, G. and Carbon, J. (1980) Gene 10, 157-166
6 Hartley, J.L. and Donelson, J.E. (1980) Nature 286, 860-864
7 DeFilippes, F.M. (1974) Biochem. Biophys. Res. Commun. 58, 586-596
8 Bickle, T.A., Pirrotta, V., and Imber, R. (1977) Nucl. Acids Res. 4, 2561-2572
9 The clones were grown under L2/B1 conditions according to the German "Richtlinien zum Schutz vor Gefahren durch in vitro-neukombinierte DNA"
10 Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523
11 Southern, E.M. (1975) J. Mol. Biol. 98, 503-517
12 Lillehaug, J.R., Kleppe, R.K., and Kleppe, K. (1976) Biochemistry 15, 1858-1865
13 Jeffreys, A.J. and Flavell, R.A. (1977) Cell 12, 429-439
14 Maxam, A.M. and Gilbert, W. (1980) Methods Enzym. 65, 499-560
15 Gray, C.P., Sommer, R., Polke, C., Beck, E., and Schaller, H. (1978) Proc. Natl. Acad. Sci. USA 75, 50-53
16 Feldmann, H. (1976) Nucl. Acids Res. 3, 2379-2386
17 Neumaier, P., Dissertation München 1981
18 Kobayashi, T., Irie, T., Yoshida, M., Takeishi, K., and Ukita, T. (1974) Biochim. Biophys. Acta 366, 168-181
19 Sprinzl, M., Grüter, F., and Gauss, D.H. (1978) Nucl. Acids Res. 5, r15-r28
20 Struhl, K., Stinchcomb, D.T., Scherer, S., and Davis, R.W. (1979) Proc. Natl. Acad. Sci. USA 76, 1035-1039
21 Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979) Nature 282, 39-43
22 e.g.: Zakian, V.A., Brewer, B., and Fangman, W.L. (1979) Cell 17, 923-934
23 Crews, S., Ojala, D., Posakony, J., Nishiguchi, J., and Attardi, G. (1979) Nature 277, 192-198
24 Benz, E.W., Reinberg, D., Vicuna, R., and Hurwitz, J. (1980) J. Biol. Chem. 255, 1096-1106
25 Shortle, D. and Nathans, D. (1979) J.Mol.Biol. 131, 801-817

26 Tinoco, I.,jun., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M., and Gralla, J. (1973) Nature new Biol. 246, 40-41
27 Zasloff, M. and Santos, T. (1980) Proc. Natl. Acad. Sci. USA 77, 5668-5672
28 Chan, C.S.M. and Tye, Bik-Kwoon (1980) Proc. Natl. Acad. Sci. USA 77, 6329-6333