Volume 10 Number 21 1982 Nucleic Acids Research

Identification of splicing signals in introns of yeast mitochondrial split genes: mutational alterations in intron bI1 and secondary structures in related introns

Carlo Schmelzer, Cornelia Schmidt and Rudolf J.Schweyen

Institut für Genetik und Mikrobiologie, Universität München, Maria-Ward-Str. 1a, D-8000 München 19, FRG

Received 20 July 1982; Revised and Accepted 1 October 1982

ABSTRACT

Four mitochondrial mutations are known to block excision of intron I1 of the cob gene in S.cerevisiae. The nucleotide sequence alteration of one of them, M4873, has been determined. It is a deletion of 1 bp in a run of five G's at a distance of 30 to 34 bp upstream to the 3' splice point. Reversion is found to occur by restoration of the run of five G's either by insertion of 1 G (wild type reversion) or by transition A→G next to this run of G's (pseudo-wild type reversion). The effect of mutation and reversion on RNA splicing indicates that the run of five G's is of critical importance for intron I1 excision, possibly in participating in the formation of a splice signal with a helical structure. This presumption is confirmed by the observation that this sequence is part of a larger sequence of some 80 bp next to the 3' splice point which is conserved to some extend in the four mitochondrial introns (bI1, aI1, aI2, aI5) that survive after excision as circular RNAs. Most striking is the conservation of this sequence at the level of secondary structure.

INTRODUCTION

In yeast mitochondria three genes have been found to be split: cob, coding for apocytochrome b, oxi 3, coding for subunit I of cytochrome oxidase and the gene coding for the large ribosomal RNA (for review see 1). The nucleotide sequences of the introns published so far reveal two interesting features, discriminating mitochondrial introns from those in nuclear genes:

(i) Long open reading frames (orfs), mostly in phase with preceding exons, are frequent. Genetic evidence suggests that two of these orfs code for proteins - the so-called maturases - with so far unknown functions in excision of the respective introns,

0305-1048/82/1021-6797$ 2.00/0

(ii) nucleotide sequences next to the splice points are variable such that no consensus sequence could be established for 5'- and 3' ends of mitochondrial introns. Moreover, no other nucleotide sequence of any meaningful length appears to be in common for all mitochondrial introns.

There are mutations in introns which - according to genetic evidence - block the excision of the respective intron by altering splicing signal sequences but leaving the intron encoded maturases intact. Some of these mutational alterations have been determined by nucleotide sequence analysis. Whereas one mutation was found to affect the 5' exon-intron boundary, several others were shown to alter sequences inside the intron, either at some distance to the 5' end or to the 3' end (2, 3, 4). Thus, internal sequences of introns apparently are involved in RNA splicing in mitochondria. This contrasts with the situation in nuclear genes, where the exon-intron boundaries only are found to be of critical importance (for review see 5).

In this communication the approach of identifying splicing signals by sequencing intron mutations is applied to intron I1 of the cob gene (bI1). This is one of the introns which do not code for a protein and whose excision is independent of mitochondrial translation and its products (6, 7, 8).

We report a primary mutational alteration in a nucleotide sequence at some distance upstream from the 3' end of bI1 , resulting in a block of bI1 excision. Reversion to pseudo-wild type is found to be possible by a secondary mutation in the same sequence. The nucleotide sequence at the 3' end of bI1 is found to be conserved in other introns. This supports the idea that it is of critical importance for RNA splicing.

MATERIALS AND METHODS

Yeast strains: The mutants studied are derived from strain 777-3A (long cob gene) (9). Revertants of M4873 were either spontaneous or isolated after treatment with $MnCl_2$. Temperature-dependent growth was tested on nonfermentable substrates at $20^{\circ}C$, $28^{\circ}C$, and $35^{\circ}C$.

Preparation of mtDNA: Yeast cells were grown in a medium con-

taining 5% glucose and 1% yeast extract and harvested in the early stationary phase. Mitochondria were isolated from protoplasts prepared by the use of Zymolase 60000 (Miles, Germany). The mtDNA was purified by centrifugation in CsCl gradients with bisbenzimide dye (Hoechst 33258).

DNA-sequencing: mtDNA was digested with restriction endonucleases (Bethesda Research Lab.) according to the recommendations of the producer. A 4300 bp HpaII restriction fragment, which extends from the 5' end of the cob region to bI3 was digested with BglII. The BglII ends were selectively labeled with DNA polymerase 'Klenov Fragment' (Boehringer Mannheim) using ^{32}P-dGTP (2000-3000 Ci/mMol, NEN). Labeled fragments were separated on 1.5% agarose gels and both fragments were sequenced according to the procedure described by Maxam and Gilbert (10). Gel runs allowed the determination of 150-200 nucleotides upstream and downstream from the BglII site next to the 3' splice point of intron I1 (cf. Fig. 1a).

Gel electrophoresis of mtRNA: Aliquots of 2 µg mtRNA were separated by electrophoresis on 1.5% agarose slab gels containing 5 mM methyl mercury hydroxide for 4 hours at room temperature with an applied voltage of 160 V (11). S-values of transcripts and their corresponding sizes in bases were determined by calibration with 15S mitochondrial rRNA (1600 bases) and 21S mitochondrial rRNA (3200 bases) (12,13).

RNA immobilization and hybridization with ^{32}P-labeled mtDNA: Preparation of diazobenzyloxymethyl (DBM) paper and gels for RNA transfer, prehybridization and hybridization were carried out as described by Alwine et al. (14). ^{32}P-labeled mtDNA which covers the entire cob region was prepared by nick-translation (15) and used as cob specific probe for hybridization.

RESULTS AND DISCUSSION

Among the many cytochrome b deficient mutants (cob$^-$) four are known to prevent expression of the cob gene by blocking the excision of

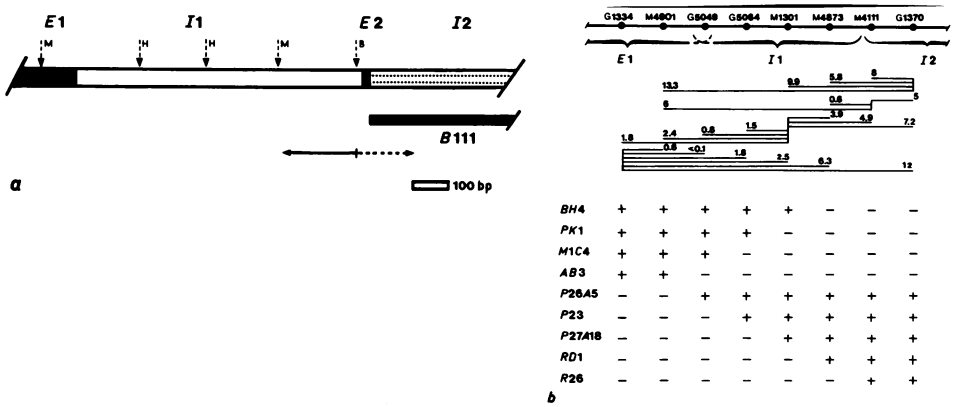

Fig. 1a Physical map of the 5' part of the long form of the apocytochrome b gene (cob) including the first intron (I1) and adjacent exons. Solid bars, exons; open bars, introns. The dotted bar denotes an open reading frame which is in phase with upstream exon E2 and is assumed to code for a maturase (7). Restriction sites for MboI (M), HinfI (H), and BglII (B) are given on top of the figure. The heavy black bar represents the mtDNA sequence in rho⁻ B111 according to the data from Lazowska et al. (7); this rho⁻ clone does not restore wild phenotype in crosses with the I1 mutant M4873. Horizontal arrows designate the sequenced region.

Fig. 1b Genetic localization of mutants in E1 and I1 of the cob region by recombination analysis (upper part) and rho⁻ deletion mapping (lower part). Horizontal lines in the upper part designate crosses between two different cob⁻ mutants; the percentage of respiring recombinants is given above the line. The cob⁻ mutants with prefix G originate from Gif (22), those with prefix M from the Munich collection (9, 23). In the lower part (+) indicates that a given rho⁻ clone (left side of the figure) has retained the wild type allele of a given cob⁻ mutation, whereas (-) indicates that a rho⁻ clone has lost the corresponding alleles. PK1, P26, P23, P27, RD1, and R26 were characterized by P. P. Slonimski, S. Kotylak, P. Perlman, and R. J. Schweyen (unpubl.), M1C4 was isolated by Bechmann et al. (9), P26A5 and P27A8 are subclones of P26 and P27, respectively.

intron I1 from cob pre-mRNA and - as a consequence of this - the excision of downstream introns (16, 17). Genetic mapping by rho⁻ deletion analysis places these four mutations between presumable exon E1 mutations and intron I2 mutations (Fig. 1b). Recombinant frequencies observed in crosses of these four mutants with exon E1 and intron I2 mutations confirm this map position. The data suggest that (i) G5049

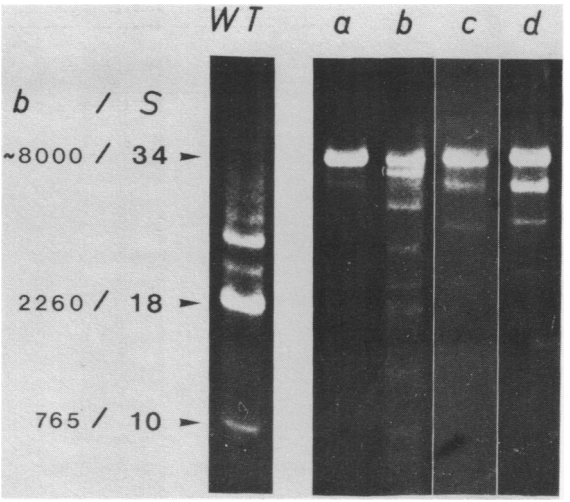

Fig. 2 Comparison of cob transcripts from wild type and I1 mutants. mtRNA was separated by electrophoresis on 1.5% agarose gels containing methyl mercury hydroxide according to Bailey and Davidson (11), transferred to DBM paper (14) and hybridized with ^{32}P-labeled mtDNA of the cob region. WT, wild type; a: G5049, b: G5064, c: M1301, d: M4873.

and G5064 map next to exon E1 mutations, possibly in the exon E1 intron I1 boundary, (ii) M1301 maps at a short distance to exon E1 mutations, probably in the 5' part of I1, and (iii) M4873 maps close to the most upstream intron I2 mutation M4111, possibly next to the 3' end of intron I1.

These four mutants accumulate a major transcript of ~ 8000 bases (Fig. 2; 34S); this RNA has been previously identified as the largest cob pre-mRNA containing the intron I1 sequence as well as the other intron sequences (17). Smaller transcripts are detected as minor species in G5049 and M1301 (30S) or M4873 (30S, 26S, and 20S). They have been shown to contain intron I1 but to lack one or several of the other introns. G5064 also exhibits a series of minor transcripts; their sizes are different from those in the other mutants. Since their composition has not been determined so far, excision of intron I1 in a minor fraction of the cob pre-mRNA cannot be excluded. However, the major effect, the accumulation of 34S RNA and the map position characterize G5064 as an I1 mutant.

Intron I1 does not code for a protein. Its excision occurs in the absence of mitochondrial protein synthesis (8) and even in rho$^-$ cells lacking a mitochondrial translation apparatus (6). Thus, the mutations in intron I1 or in its boundaries , which block I1 excision, most likely

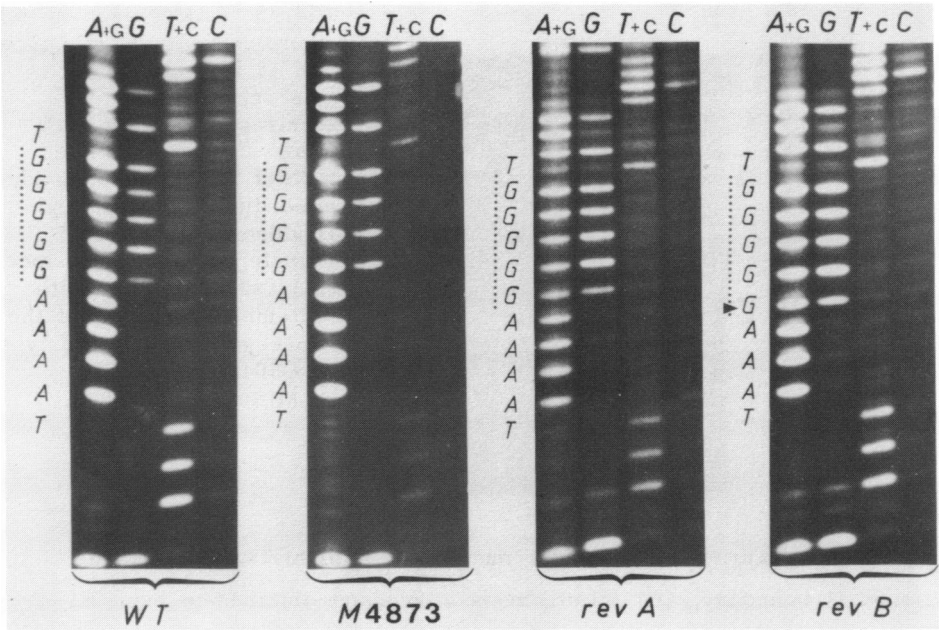

Fig. 3 Sequencing gels demonstrating the alterations in mutant M4873 and in two types of revertants from this mutant (revA and revB). The sequence of a fragment (1600 bp HpaII/BglII) 3'end-labeled at the BglII site is shown on a 20% gel. This fragment extends from the 3' end of I1 into the 5' direction (cf. Fig. 1a). The dotted line denotes the sequence in which the deletion of 1 base occured (M4873). In revertants of type A the wild type sequence is restored, whereas in type B revertants a A → G transition has occured, resulting in an A deletion. DNA sequences shown correspond to the noncoding strand.

alter a nucleotide sequence which per se serves as a splicing signal.

We set out to identify one of these signals by sequencing mutation M4873. Since genetic mapping suggests a location next to the 3' splice point, we focussed our attention on sequences of about 200 bp upstream and downstream from the BglII site next to the presumed 3' splice point as given by Lazowska et al. (7). The only mutation detected in this sequence is a - 1 base deletion in a run of five G's located at a distance of 30 to 34 bp from the presumed splice point. Figure 3 displays part of the sequencing gels showing the run of five G's (wild type) and four G's (M4873).

In order to demonstrate that this deletion is not a silent mutation

but causes the splicing defect in M4873 we have investigated the same region in eight revertants of M4873; (revertants of M4873 occur spontaneously at a frequency of about 10^{-6}). All eight revertants reveal the restoration of the run of five G's. In five of them this is achieved by a + 1G insertion; they are true wild type revertants (Fig. 3, revA). In the other three revertants the run of five G's is restored by a transition of an adjacent A to G, resulting in a - 1A deletion in a run of four A's (Fig. 3, revB). These are pseudo-wild type revertants. Their phenotype differs from true wild type revertants in that cell growth on nonfermentable substrates is slightly reduced at 35°C.

We deduce from these data that the run of five G's at a distance of 30 to 34 b from the presumed 3' splice point is of critical importance for the excision of I1 from the pre-mRNA.

The mutated sequence is part of an open reading frame of 37 triplets in the 3' part of I1. Since it is in frame with the downstream orf exon E2 - intron I2 it might be part of the I2 (box 3) maturase coding sequence as discussed by Lazowska et al. (7). However, our data render this possibility unlikely: (i) M4873 specifically blocks excision of I1 (17), not of I2 as do mutations in the box 3 cluster of I2 (17,18), (ii) type B reversions leave the reading frame blocked at the same position as does M4873, but restore wild type functions almost completely. Thus we have not selected the restoration of the open reading frame but the restoration of a nucleotide sequence, the run of five G's, which per se serves as a signal in RNA splicing or which is part of a larger signal sequence.

A mutation at a comparable distance to the 3' end of an intron has only been found in cob I4 (2), but neither the sequence altered nor the surrounding sequences in I1 and I4 show homology. A homologous counterpart to the cob I1 3' end, however, is found in intron I5 of the oxi 3 gene (aI5; nomenclature and sequence as in ref. 19). As shown in Fig. 4a, a stretch of 80 bases upstream from the presumed splice points appears to be highly conserved. The sequence which in cob I1 carries the mutation M4873 (GGGGGAAAA) is perfectly conserved and has the same position relative to the 3' end in cob I1 and oxi 3 I5. A

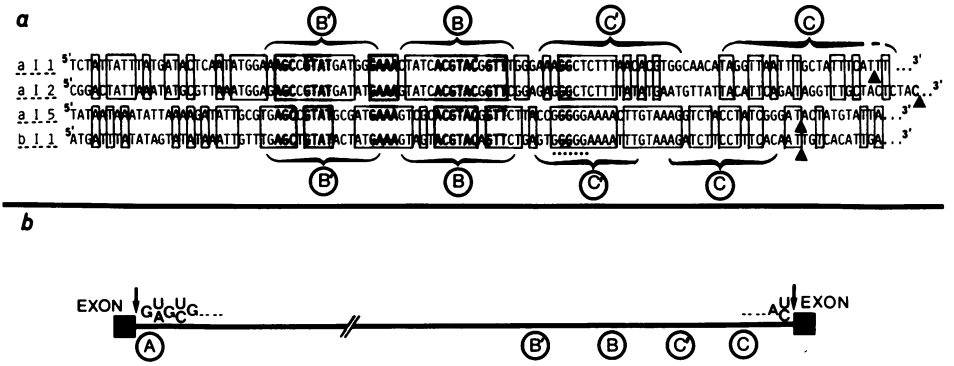

Fig. 4a Comparison of DNA sequences of the 3' part of introns from the cob region (bI1) and the oxi 3 region (aI1, aI2, aI5). Sequences are according to Lazowska et al. (7) and Bonitz et al. (19). Open boxes designate sequences homologous either within introns aI1 and aI2 or within aI5 and bI1. Shaded areas denote regions which are conserved in all four introns. Exon-intron boundaries are marked by arrows. The dotted line marks the sequence altered in M4873. Letters with circles denote regions which could serve as signals for splicing by building characteristical secondary structures (see text).
Fig. 4b Features common to introns bI1, aI1, aI2, and aI5. (For details see text).

major part of the sequence conserved in these introns is also detected in aI1 and aI2, although somewhat distant to the presumed splice point (Fig. 4a). The signal sequence GGGGGAAAA is not conserved except for the 5' $\begin{matrix}G\\A\end{matrix}$GG, but it is worth noting that both aI1 and aI2 replace it by the same octanucleotide (GGCTCTTT). Conserved are further a pentanucleotide immediately after the 5' splice point and a dinucleotide in front of the 3' splice point (Fig. 4b). Neither these nor the sequence homology in the 3' part are detected in any of the other known mitochondrial introns. There is one more result which distinguishes bI1, aI5, aI1, aI2 as having splicing related criteria of their own: The excised introns survive as populations of covalently closed circles with sizes of 10S (bI1), 11S (aI5), 19S (aI2), and 19.2S (aI1) (6, 20, 21). In classifying mitochondrial introns we may thus regard these four as members of one family having similar pentanucleotide sequences at the 5' end and a long sequence at the 3' end. The internal parts seem to

be variable, aI1 and aI2 carry an open reading frame in phase with the preceding exon as do introns not belonging to this family.

What is the function of the conserved intron sequences ? The considerable length of the conserved sequence, that is 80 bases in introns bI1 and aI5, suggests that they are involved in secondary structures rather than that only the primary sequence is recognized as a splicing signal. Analysis of the highly conserved sequence carrying mutation M4873 allows us to postulate the presence of a helical structure ending in front of the presumed 3' splice point in bI1 and aI5. The idea that this structure has its biological significance as a splice signal (or part of it) finds strong support in the following observations:

(i) The helical structures in bI1 and aI5 are identical, having a first stem region with a bulge loop (A or U) followed by one or three mismatches and a second stem ending in a loop (GUAA) (Fig. 5); 8 of 33 bases differ in bI1 and aI5, but 6 of the exchanges are compensatory, replacing one pair of bases by another such that the stem and loop structure is conserved; in other words: maintenance of secondary structure has precedence over a conservation of the primary sequence.

(ii) Sequences directly preceding the 3' splice point in aI1 and aI2 can form a similar stem and loop structure with a strong conservation of the section containing the bulge loop, although primary sequences vary (Fig. 6).

(iii) Mutation M4873 (- 1G deletion) alters this structure in bI1 by increasing the bulge loop ($\overset{\cdot\cdot UA^{A}AAGG\ GGU\cdot\cdot}{\cdot\cdot AU_{C}UUCC_{UU}UCA\cdot\cdot}$, cf. Fig. 5) and decreasing the stability of the stem. This alteration might be sufficient to diminish the recognition of the hybrid as a splice signal. Two reversions of M4873 are found: Addition of 1G, restoring the wild type sequence, and transition A → G next to the run of G's which does not restore the WT sequence but the secondary structure ($\overset{\cdot\cdot UA\ AAGGG\ GGU\cdot\cdot}{\cdot\cdot AU_{C}UUCC_{U}UCA\cdot\cdot}$). Among eight revertants five show the + 1G addition and three the A → G transition; it thus appears, that restoration of the splice signal is achieved only by mutations restoring the hybrid structure with a bulge loop as shown in Fig. 6.

There is another sequence further upstream (Fig. 4, Ⓑ Ⓑ')

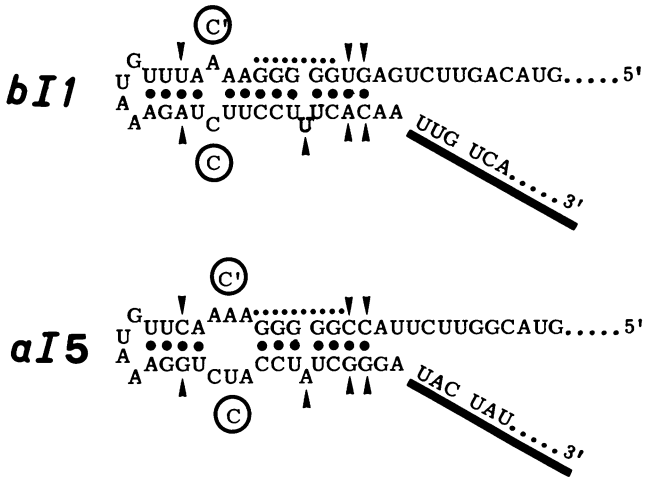

Fig. 5 Possible interactions of intron boundaries within aI5 and bI1. Possible RNA secondary structures within the 3' parts of introns bI1 and aI3 are shown. Exons are marked by heavy lines; dotted line: Sequence altered in M4873. The arrows either indicate base changes within the bulge loop or compensatory changes of base pairs resulting in conservation of the secondary structure within this part of the two introns. Letters in circles correspond to those shown in Fig. 4.

which is conserved in all four introns (bI1, aI5, aI1, aI2) and which can also form a quite stable helical structure (not shown). We therefore assume that the structure Ⓑ Ⓑ' may be also part of a splicing signal. A computer analysis of mitochondrial intron sequences by

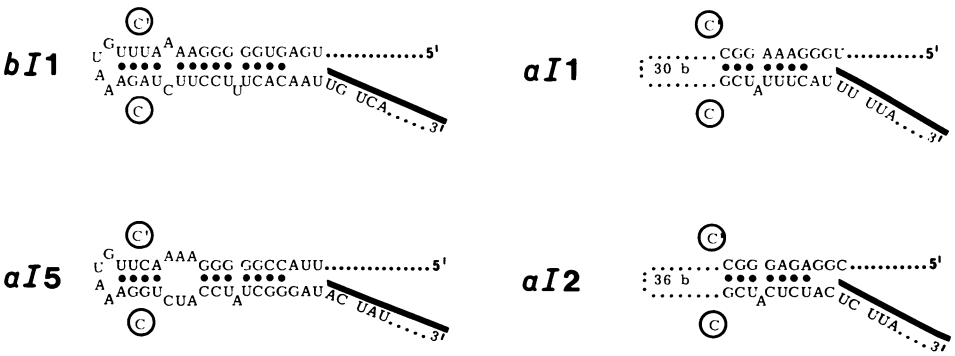

Fig. 6 Comparison of conserved RNA secondary structures within the 3' part of bI1, aI1, aI2 and aI5. (Signs as in Fig. 5.)

Michel et al. (24) confirms that both hairpins in the 3' part, Ⓑ Ⓑ' and Ⓒ Ⓒ' , are energetically favorable hybrids and that they are conserved.

To summarize, the data presented above yield first evidences that in certain introns of yeast mitochondria conserved secondary structures at the 3' end of introns are important for their excision. We could not detect conserved helical areas at the 5' end of these introns; however, there is a primary sequence common to all four introns ($\downarrow G^{U}_{A}G^{U}_{C}G\ldots$, Fig. 4b Ⓐ). To test the importance of this sequence (e.g. its possible interaction with sequences at the 3' end of introns by base pairing) we intend to analyze further mutational alterations in the four introns bI1, aI1, aI2 and aI5.

ACKNOWLEDGEMENTS

We thank B. Dujon for sending us a preprint of his work prior to publication, M. Dobres and A. Haid for helpful discussions.

This work was supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Dujon, B. (1981) In Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance. Strathern, J.N., Jones, E.W., and Broach, J.R., eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), monograph 11, 505-635.
2. Jacq, C., Pajot, P., Lazowska, J., Dujardin, G., Claisse, M., Groudinsky, O., de la Salle, H., Grandchamp, C., Labouesse, M., Gargouri, A., Guiard, B., Spyridakis, A., Dreyfus, M., and Slonimski, P.P. (1982) In Mitochondrial Genes. Slonimski, P.P., Borst, P., and Attardi, G., eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), monograph 12, in press.
3. Weiss-Brummer, B., Rödel, G., Schweyen, R.J., and Kaudewitz, F. (1982) Cell 29, in press.
4. De La Salle, H., Jacq, C., and Slonimski, P.P. (1982) Cell 28, 721-732.
5. Lewin, B. (1980) Cell 22, 324-326.
6. Halbreich, A., Pajot, P., Foucher, M., Grandchamp, C., and Slonimski, P.P. (1980) Cell 19, 321-329.
7. Lazowska, J., Jacq, C., and Slonimski, P.P. (1980) Cell 22, 333-348.
8. Schmelzer, C., and Schweyen, R.J. (1982) Nucleic Acids Res. 10, 513-524.

9. Bechmann, H., Haid, A., Schweyen, R. J., Mathews, S., and Kaudewitz, F. (1981) J. Biol. Chem. 256, 3525-3531.
10. Maxam, M. A., and Gilbert, W. (1980) In Methods in Enzymology 65, Moldave, K., and Grossman, L., eds. (New York: Academic Press), 499-560.
11. Bailey, J. M., and Davidson, N. (1976) Anal. Biochem. 70, 75-85.
12. Heyting, C., Meijlink, F. C. P. W., Verbeet, M. Ph., Sanders, J. P. M., Bos, J. L., and Borst, P. (1979) Mol. Gen. Genet. 168, 231-250.
13. Tabak, H. F., Hecht, N. B., Menke, H. H., and Hollenberg, C. P. (1979) Curr. Genet. 1, 33-43.
14. Alwine, J. C., Kemp, D. J., and Stark, G. R. (1977) Proc. Nat. Acad. Sci. USA 74, 5350-5354.
15. Rigby, P. J. W., Dieckman, M., Rhodes, C., and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
16. Haid, A., Grosch, G., Schmelzer, C., Schweyen, R. J., and Kaudewitz, F. (1980) Curr. Genet. 1, 155-161.
17. Schmelzer, C., Haid, A., Grosch, G., Schweyen, R. J., and Kaudewitz, F. (1981) J. Biol. Chem. 256, 7610-7619.
18. Van Ommen, G. J. B., Boer, P. H., Groot, G. S. P., de Haan, M., Roosendaal, E., Grivell, L. A., Haid, A., and Schweyen, R. J. (1980) Cell 20, 173-183.
19. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E., Tzagoloff, A., and Macino, G. (1980) J. Biol. Chem. 255, 11927-11941.
20. Arnberg, A. C., Van Ommen, G. J. B., Grivell, L. A., Van Bruggen, E. F. J., and Borst, P. (1980) Cell 19, 313-319.
21. Grivell, L. A., Arnberg, A. C., Hengens, L. A. M., Roosendaal, E., Van Ommen, G. J. B., and Van Bruggen, E. F. J. (1980) In The organization and expression of the mitochondrial genome. Saccone, C., and Kroon, A. M., eds. (Elsevier/North-Holland, Amsterdam), 37-49.
22. Claisse, M. L., Spyridakis, A., and Slonimski, P. P. (1977) In Mitochondria 1977: Genetics and biogenesis of mitochondria. Bandlow, W., Schweyen, R. J., Wolf, K., and Kaudewitz, F., eds. (Walter de Gruyter, Berlin), 337-344.
23. Schweyen, R. J., Weiss-Brummer, B., Backhaus, B., and Kaudewitz, F. (1978) Mol. Gen. Genet. 159, 151-160.
24. Michel, F., Jacquier, A., and Dujon, B. (1982) Biochimie, in press.