# 5 Modulation of Replication Efficiency by Non-coding Sequences and by the Environment

**GIORGIO BERNARDI, REGINA GOURSOT, EDDA RAYKO, RENÉ GOURSOT, BAYA CHERIF-ZAHAR, AND ROBERTA MELIS**
*Laboratoire de Génétique Moléculaire, Institut Jacques Monod, Paris, France*

## 5.1 INTRODUCTION

Two recent observations from our laboratory are relevant to the subject of this volume. They both concern the modulation of replication of the mitochondrial genome from cytoplasmic petite mutants of yeast. In the first case, modulation is due to non-coding sequences flanking the *ori* sequences, in the second to an environmental factor, temperature.

The underlying questions are of a very fundamental nature. In the first case, the problem is that of the influence of sequences flanking a given DNA segment on the functional properties of the latter. Obvious effects can be expected, and are found, owing to the presence or absence of regulatory signals in the flanking sequences. A more interesting situation arises when this is not the case and the effects are presumably just due to changes in DNA structure around the sequence under consideration. This is exactly the case investigated by us, by looking at the influence of non-coding sequences that flank the *ori* sequences of the mitochondrial genome of yeast.

The second question concerns the effect of the environment on the genome. In evolution, the environment mainly acts as a selection agent; organisms adapt to diverse environments by being selected for. The environment can also cause, in a direct or an indirect way, chemical alterations in the genome, as exemplified by the actions of mutagens. Another way in which the environment can affect the genome is by directly

*Introduction of Genetically Modified Organisms into the Environment.* Edited by Harold A. Mooney and Giorgio Bernardi
© 1990 SCOPE Published by John Wiley & Sons Ltd

inducing changes in its higher-order structure, as shown in the results to be described here.

## 5.2 THE MITOCHONDRIAL GENOME OF YEAST

The mitochondrial genome of wild-type, respiratory competent *Saccharomyces cerevisiae* is made up of 25–50 genome units (in haploid and diploid cells, respectively), which segregate into the buds at cell replication. Each one of the identical genome units (Figure 5.1) comprises genes encoding ribosomal RNAs and transfer RNAs, as well as genes encoding protein subunits of enzyme complexes associated with respiration and energy production. The latter genes contain, in a number of strains, intervening sequences, some of which contain open reading frames. All coding sequences only represent one third of the mitochondrial genome, the rest being formed by non-coding, mostly intergenic, sequences. The latter are formed by long AT spacers, only 5% in GC, and by about 200 GC clusters, about 60% in GC (de Zamaroczy and Bernardi, 1985, 1986a, 1986b, 1987).

As intergenic sequences are highly repetitive, they confer an extreme instability to the mitochondrial genome of yeast, which is dispensable (yeast being able to survive through fermentation). Indeed, excisions of DNA segments from the mitochondrial genome occur very frequently at pairs of direct repeats essentially located in intergenic sequences (de Zamaroczy *et al.*, 1983). The excised segment is amplified in tandem to become the repeat unit of the defective genome of the cytoplasmic respiratory deficient petite mutants. These defective genomes segregate into the buds and may undergo secondary excisions (Figure 5.2). Replication of spontaneous petite mutants is usually insured by the presence in its repeat units of one of the three to four functional *ori* sequences (de Zamaroczy *et al.*, 1984; Baldacci *et al.*, 1984) from the wild-type genome (Figure 5.3). Such *ori* sequences are about 300 bp long and comprise several features : (1) two GC clusters, A and B, which can form with the intervening AT sequences an A–B fold (see Figure 5.5); (2) a long AT sequence *l*; and (3) a GC cluster, C, which is followed by a region *r* in which mitochondrial RNA polymerase starts the synthesis of the RNA primer of one nascent DNA chain; the RNA primer for the other nascent DNA chain is started at the other side of cluster C (Figure 5.3).

## 5.3 THE EFFECT OF DNA REGIONS FLANKING THE ORI SEQUENCE ON REPLICATION EFFICIENCY

If one tests the replicative efficiency of *ori+* petites (namely of petites carrying an intact *ori* sequence in their mitochondrial genome) by crossing

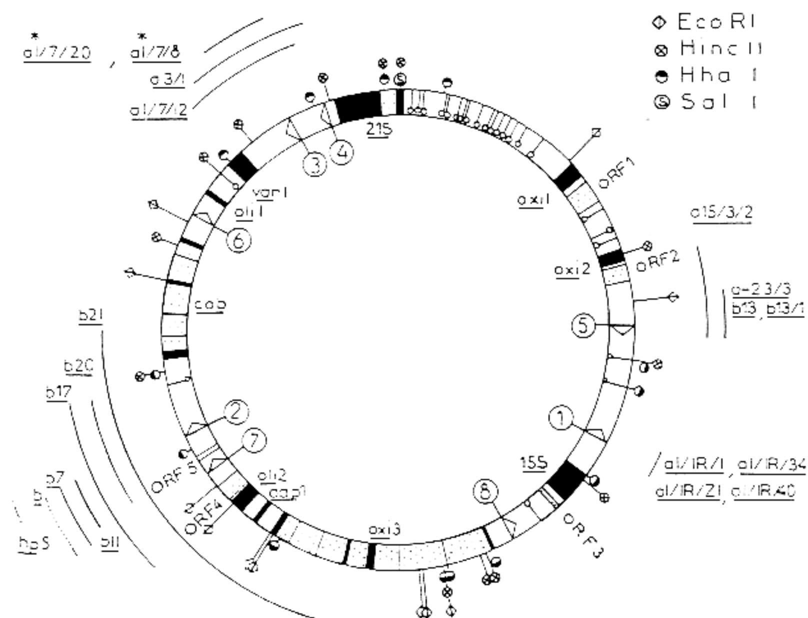


**Figure 5.1** Physical map of a long mitochondrial genome unit of wild-type *S. cerevisiae*. Black areas correspond to mitochondrial genes or their exons (the last exon block of *oxi*3 comprises exons A 6–8 and introns a1 7,8); dotted areas to intervening sequences and intergenic open reading frames (ORF1–ORF5; Colin *et al.*, 1985); radial lines indicate tRNA genes (the thr1 gene is the only one to have an anticlockwise orientation). White areas correspond to long AT spacers with embedded short GC clusters. Among mitochondrial genes, *oxi*1, 2, and 3 encode subunits II, III, and I, respectively, of cytochrome c oxidase; *cob*, cytochrome b; *aap*1, *oli*2, and *oli*1, subunits 8,6, and 9 of ATPase; *var*1, a protein associated with the small mitochondrial ribosome subunit; 9S corresponds to the central part of tRNA synthesis locus; 15S and 21S are the genes for the small and large ribosomal RNAs, respectively. Triangles indicate the location of *ori* sequences 1–8; they point in the direction cluster C to cluster A.

The localization of the repeat units from the mitochondrial genomes of petite mutants used in the present work on the physical map of the mitochondrial genome of wt cells is shown. When several petite genomes correspond to the same localization, they are of about equal sizes. In the case of petite b, the dotted line corresponds to a central deletion; in that of petite hp5, excision sequences were not determined. Some restriction sites are indicated.
(From Rayko *et al.*, 1988.)

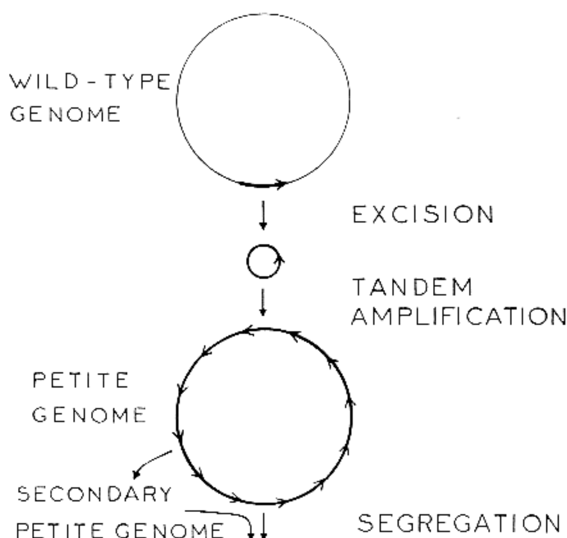


**Figure 5.2** Scheme of the mechanism of formation of the defective mitochondrial genomes from spontaneous petite mutants. A mitochondrial genome unit from wild-type cells undergoes an excision. The excised segment is tandemly amplified to form a mitochondrial genome unit of the petite mutant. This genome replicates and segregates into the buds. It may also undergo secondary excisions (from Bernardi, 1983).

with either a wild-type strain (suppressivity test) or another petite, the general rule is that the mitochondrial genome made up of shorter repeat units (and containing therefore more *ori* sequences) will show a higher replicative efficiency, as shown by its predominance in the progeny. We have found, however, exceptions to this rule (Figure 5.4). Indeed, some petite genomes characterized by longer repeat units may replicate better than other ones carrying shorter repeat units. Since comparisons were made among petite genomes carrying the same *ori* sequences, the only difference was the presence or absence of a few dozens of nucleotides in the regions flanking the *ori* sequences. Such regions are made up by AT spacers and GC clusters, may be located upstream or downstream of *ori* sequences, are different in length and primary structure, and therefore unlikely to carry any specific sequence signal.

The conclusion of these investigations (Rayko *et al.*, 1988) is that non-coding sequences flanking the *ori* sequences may affect the efficiency of replication in all likelihood by altering the higher-order structures of the mitochondrial genome. This conclusion appears to be of general interest, since it does not seem to be limited to mitochondrial *ori* sequences of yeast. Indeed, the deletions in a 500 bp region of plasmid PT181 (a 4.4 Kb

```
CCGAAACTAAATAAAATATATTATTAATAATATTATATAATATAATAATATATAATAATTTTATATAAATATATATTTATATATTAAATTAAATTATAA
                                                50
TTTTATTATGAAAATTATATCTTTTATTATATTCTCCTTTCGGGGTTCCGGCTCCGGCCGGAAAATCCCTCCCAGTCCTGACGGGGCTTGCTCACATCCT
                                                150                                    εZ1
TATAATTAAATAAAGGTGCCAATATATTTTATATATAAATTAATATATATATATTATATATATATATATACTTTTTTTTAATAATATATCTATATATATAAA
                                                250                    A               p
TTAATGTATTATATTATATTTTTATATAATATATTATTAATTATTATTTTAATTTTCTATTCTATTGTGGGGTCCCAATTACTATTTTCAATAATAATT
  s      B                                      350                   ℓa        ε26        ε14
ATTATTGGGACCCGGATATCTTCTTGTTTATCATTTATTATTTTATTAAATTTATTATTATTTTAATTTATATTTATATTATATAATTAATTATATCGT
                                                450                    ℓ
TTATACCTTATTATTATTATATAATATATTATATTATTATAATATATTTATTGATTATATTATAAAATTTATTCTATGTGTGTTCTATATATATTTAATA
                       C              r′             r              ra
TTCTGGTTATTGATCACCCACCCCCTCCCCCTATAAAACTTAATTTATTACTTATATATTTATAAATATAAATCTAACTTAATTAATAATTTAAATAATA
                                                650                        ε26          εZ1
TACTTAATATTTAATAAATAGGAATAATTGTAACCTTTTTATAATTATATATAATAATAATATATATTATCAAATAATTATTATTTCTTTTTTTTCTTTA
                                                750                      ε14
ATTAATAATTAATTAATTAATTAATTAATATTTTATTTTATAAAAAATATATTTCTCCTTACGGGGTTCCGGCTCCGTAGCGGG
                                                850
```

**Figure 5.3** The repeat unit of the mitochondrial genome of primary petite a1/1R/1 (see Gaillard *et al.*, 1980) is 884 bp in size and contains an *ori* sequence, *ori*1. This comprises three GC clusters, A, B and C (heavy overlines). Cluster B is separated from cluster A by two short AT sequences *p* and *s*, and from cluster C by a long AT sequence *l* (broken overline). Cluster C is followed by sequence *r* (light overline) which comprises, on the strand shown, a nonanucleotide (heavy underline) where transcription of the RNA primer of one nascent DNA chain starts (arrow pointing to the left); transcription of the RNA primer for the other nascent DNA chain starts before cluster C (arrow pointing to the right) (Baldacci *et al.*, 1984). Sequence *r* is followed by a sequence *ra* which might pair with sequence *la* (within the *l* sequence) to form a hypothetical tertiary structure (de Zamaroczy *et al.*, 1984). In the case of *ori*1, sequence *r* is preceded by a duplication, sequence *r′*. Excision sequences (ε) of the secondary mitochondrial genomes from petites 14, 26, and Z1 (see Gaillard *et al.*, 1980; de Zamaroczy *et al.*, 1984) are indicated by boxes (From Goursot *et al.*, 1988).

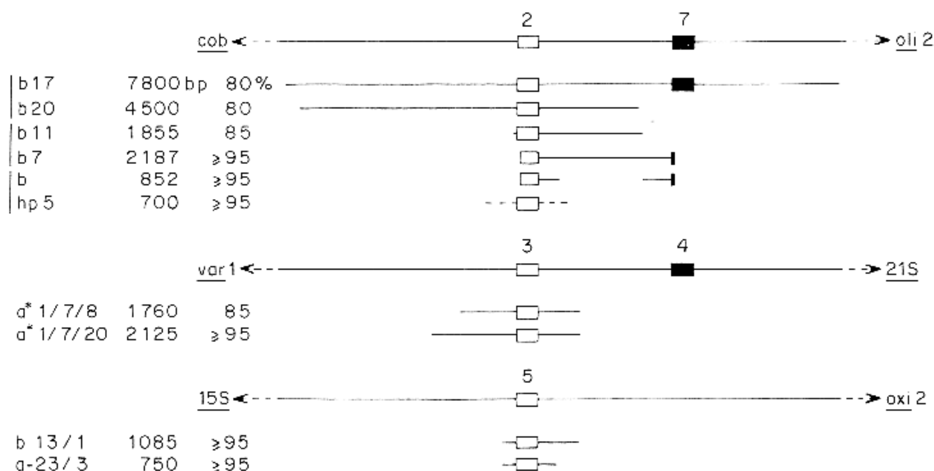


**Figure 5.4** Repeat units of petites harbouring mitichondrial genomes carrying identical *ori* sequences and exhibiting suppressivities and/or competitive abilities in petite × petite crosses which do not follow the repeat unit size rule (see Discussion). Repeat unit sizes and suppressivities are indicated. Open boxes indicate active *ori* sequences 2,3, and 5, blackened boxes indicate inactive *ori* sequences 4 and 7 (de Zamaroczy *et al*, 1984). The *ori* sequences are all oriented with cluster C to the right of cluster A. Flanking genetic markers are indicated (see also Figure 5.1). Vertical bars on the left indicate the pairs of petite repeat units which are compared and discussed. (From Rayko *et al.*, 1988.)

multicopy plasmid of *Staphylococcus aureus*) external to the minimal replicon decreased the ability of the plasmid to compete with a coexisting incompatible plasmid, whereas in the homoplasmic state the deletions affected neither copy number nor plasmid stability (Gennaro and Novick, 1986). These deletions appeared to affect the interaction of RepC, a trans-acting initiator protein which is rate limiting for replication, and the plasmid origin of replication.

Concerning the mechanism by which the extra sequences favour replication, it should be pointed out that in different cases the extra sequences are different in primary structure, size, and position relative to the *ori* sequences (Figure 5.4). This makes it unlikely that replication is favoured by specific 'replication enhancement' sequences, and rather suggests that the effect is due to differences in the DNA or nucleoid (see, for instance, Rickwood *et al.*, 1981) structure in the different petite genomes under consideration. Such differences can be visualized as differences in DNA bending, DNA superhelicity, and/or DNA–protein interactions which may modify the secondary and tertiary structure of the *ori* sequences themselves (see de Zamaroczy *et al.*, 1984) and, consequently, the initiation of DNA

replication; obviously effects on the elongation rate of newly synthesized DNA strands are also conceivable.

Since the flanking sequences of *ori* sequences in the petite genomes under consideration are made up of AT spacers and GC clusters, our results provide evidence that intergenic non-coding sequences play a role in the modulation of an essential genome function, such as replication, and substantiate the general idea, supported by other lines of evidence (Bernardi, 1982, 1983; Bernardi and Bernardi, 1986; de Zamaroczy and Bernardi, 1986a, 1987), that non-coding sequences do play a physiological role in genome function.

As a final remark, it should be noted that effects similar to those described here might be responsible for at least some of the cases in which recombinant plasmids carrying particular inserts fail to replicate.

## 5.4 THE EFFECT OF TEMPERATURE ON REPLICATION EFFICIENCY

The second observation concerns the effects of temperature on the efficiency of replication in petite genomes that have a thermosensitive structure in their *ori* sequences. Two petites from our collection, a1/1R/14 and a1/1R/26 (called 14 and 26 henceforth), are made up of repeat units containing *ori* sequences which are partially deleted in their cluster A and also in the neighbouring sequences. As a consequence, these *ori* sequences lack the A–B fold which requires an intact A cluster. They can form, however, replacement folds which are only made by AT base pairs (Figure 5.5). The lower thermal stability of the replacement fold compared to the A–B fold suggests that temperatures within a physiological range for yeast may affect its structure. We have, therefore, tested the suppressivity of these petites and compared them with that of a closely related petite which comprised, however, an intact *ori* sequence, a1/1R/Z1 (or Z1). As shown in Table 5.1, the petites containing the defective *ori* sequences showed a strong temperature dependence (in the 23–33°C range) of its replication efficiency as tested by suppressivity, whereas the control petite showed essentially no change under the same conditions.

The results presented (Goursot *et al.*, 1988) indicate that an environmental factor, temperature, can reversibly affect the replicative ability of a genome by altering its secondary (and, possibly, its tertiary) structure. Indeed, (1) these changes cannot be ascribed to enzymes involved in DNA replication, as in temperature-sensitive mutants, since the petites discussed here are isonuclear and lack mitochondrial protein synthesis, like all petites; (2) the different effects of temperature on the replicative ability of petites Z1, 14, and 26 show an excellent correlation with those expected from the secondary structures of the postulated A–B fold and replacement folds (de Zamaroczy

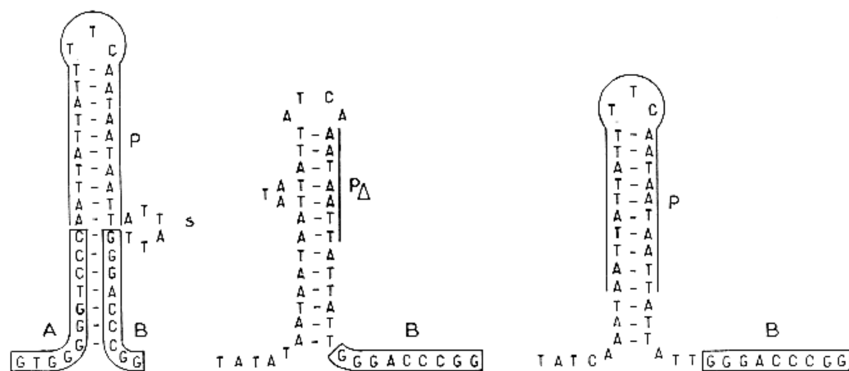


**Figure 5.5** Potential secondary structure of the 'A–B fold' of the *ori* sequence present in the repeat units of mitochondrial genomes from petites a1/1R/1 and Z1 (*ori*1) and of the 'replacement folds' that can be formed in the *ori*1 sequence present in the mitochondrial genomes of petites 14 and 26. In the case of petite 14, the residual nucleotides from the partially deleted *p* stretch (*p*Δ) can generate a hairpin structure within nucleotides from the preceding repeat unit, but the stem, only formed by 13 AT nucleotides, carries different terminal and side loops compared to the A–B fold. In the case of petite 26, the upper part of the stem and terminal loop are identical to those of the A–B fold, but the lower part is replaced by three A:T pairs (three nucleotides are derived from the preceding repeat unit). The sequences involved in the structure shown (GC clusters A and B, sequences *p* and *s*) are those indicated in Figure 5.1. (From Goursot *et al.*, 1988.)

**Table 5.1** Suppressivity of petites 14, 26, and Z1 at different temperatures

| Temperature (°C)[a] | Suppressivity of petites (%)[b] | | |
|---|---|---|---|
| | 14 | 26 | Z1 |
| 23 | 90 | 89 | 99 |
| 28 | 79 | 85 | 99 |
| 33 | 45 | 58 | 97 |

[a]Temperatures indicated were applied to precultures and to cultures used in crossings and in the incubation of diploids. Precultures in stationary phase were diluted ten times upon setting up the cultures used in crossings.

[b]Suppressivity was tested by crossing petite strains (MATa, adel) against the grand strain B (MATa, trpl, hisl). The percentage of petite/total diploid colonies was determined by counting about 1000 colonies on minimal agar.

*et al.*, 1984), an effect on tertiary structures being also possible.

Interestingly, the conclusion that DNA secondary structure is required for *ori* activity *in vivo* has also been very recently reached for bacteriophage G4, where a strong temperature-dependent impairment of replication was found after introducing by site-directed mutagenesis point mutations that

destabilize intrastrand base-pairing in the *ori* sequence (Lambert *et al.*, 1987). In the latter case, hairpin formation concerns single-stranded and not double-stranded DNA. The possibility should be left open, therefore, that hairpin formation in the petite genomes discussed here may be affected in the single-stranded DNA made during DNA polymerase action. This would not change, however, the basic conclusion that temperature differentially and reversibly affects secondary DNA structures as present during the replication of the mitochondrial genomes of petites 14 and 26.

From a general viewpoint, these results indicate the existence of a novel type of environment–genome interaction, in which reversible changes in higher-order DNA structures are induced with profound consequences on a basic genome function, such as replication. These changes concern genome transconformations which, although non-inheritable, can be maintained for many generations in the presence of the appropriate environmental condition. Genome transconformations can provide, therefore, strong selective advantages or disadvantages and play an important role in evolution, independently of classical mutations, which involve changes in the primary structure of DNA.

It should also be noted that genome transconformations of the type just described are likely to be found in organelle genomes from other poikilothermic organisms and also in prokaryotic genomes, which have a similar nucleoprotein organization and can replicate at widely different temperatures. Moreover, similar phenomena might (1) be induced by other environmental factors; (2) affect other genome functions (e.g. transcription); and also (3) be operative in other organisms.

## REFERENCES


Baldacci, G., Cherif-Zahar, B. and Bernardi, G. (1984) The initiation of DNA replication in the mitochondrial genome of yeast. *EMBO J.* **3**, 2115–20.

Bernardi, G. (1982) Evolutionary origin and the biological function of non-coding sequences in the mitochondrial genome of yeast. In: Slonimski, P.P., Borst, P. and Attardi, G. (Eds.) *Mitochondrial Genes*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 269–78.

Bernardi, G. (1983) Genome instability and the selfish DNA issue. *Folia Biol.* **29**, 82–92.

Bernardi, G. and Bernardi, G. (1986) Compositional constraints and genome evolution. *J. Mol. Biol.* **24**, 1–11.

Colin, Y., Baldacci, G. and Bernardi, G. (1985) A new putative gene in the mitochondrial genome of Saccharomyces cerevisiae. *Gene*, **36**, 1–13.

de Zamaroczy, M. and Bernardi, G. (1985) Sequence organization of the mitochondrial genome of yeast—a review. *Gene* **37**, 1–17.

de Zamaroczy, M. and Bernardi, G. (1986a) The GC clusters of the mitochondrial genome of yeast and their evolutionary origin. *Gene* **41**, 1–22.

de Zamaroczy, M. and Bernardi, G. (1986b) The primary structure of the mitochondrial genome of *Saccharomyces cerevisiae*—a review. *Gene* **47**, 155–77.

de Zamaroczy, M. and Bernardi, G. (1987) The AT spacers and the *var*1 genes from the mitochondrial genomes of *Saccharomyces cerevisiae* and *Torulopsis glabrata* : evolutionary origin and mechanism of formation. *Gene* **54**, 1–22.

de Zamaroczy, M., Faugeon-Fonty, G. and Bernardi, G. (1983) Excision sequences in the mitochondrial genome of yeast. *Gene* **21**, 193–202.

de Zamaroczy, M., Faugeron-Fonty, G., Baldacci, G., Goursot, R. and Bernardi, G. (1984) The *ori* sequences of the mitochondrial genome of a wild-type yeast strain : number, location, orientation and structure. *Gene* **32**, 439–57.

Gaillard, C., Strauss, F. and Bernardi, G. (1980) Excision sequences in the mitochondrial genome of yeast. *Nature* **283**, 218–20.

Gennaro, M.L. and Novick, R.R. (1986) Cmp, a *cis*-acting plasmid locus that increases interaction between replication origin and initiator protein. *J. Bacteriol.* **168**, 160–6.

Goursot, R., Goursot, R. and Bernardi, G. (1988) Temperature can reversibly modify the structure and the functional efficiency of *ori* sequences of the mitochondrial genome from yeast. *Gene*, **69**, 141–5.

Lambert, P.F., Kawashima, E. and Reznikoff, W.S. (1987) Secondary structure at the bacteriophage G4 origin of complementary strand DNA synthesis : *in vivo* requirements. *Gene* **53**, 257–64.

Rayko, E., Goursot, R., Cherif-Zahar, B., Melis, R. and Bernardi, G. (1988) Regions flanking *ori* sequènces affect the replication efficiency of the mitochondrial genome of *ori*+ petite mutants from yeast. *Gene* **63**, 213–26.

Rickwood, D., Chambers, J.A.A. and Barat, M. (1981) Isolation and preliminary characterization of DNA–protein complexes from the mitochondria of *Saccharomyces cerevisiae*. *Exp. Cell. Res.* **133**, 1–13.