J. Mol. Biol. (1977) **110**, 53–74

The Mitochondrial Genome of Wild-type Yeast Cells

VI†. Genome Organization

ARIEL PRUNELL‡ AND GIORGIO BERNARDI§

Laboratoire de Génétique Moléculaire
Institut de Biologie Moléculaire, Paris 5°, France

(*Received 6 April 1976, and in revised form 18 October 1976*)

Chromatography on hydroxyapatite of *Hae*III and *Hpa*II digests of yeast mitochondrial DNA shows that $0{\cdot}8\times10^6$ and 1×10^6 (molecular weight) of DNA per genome unit of 50×10^6, respectively, are present in the digests as short, single-stranded oligonucleotides, thus providing evidence for the clustering of both *Hae* sites and *Hpa* sites. An analysis of the double (*Hae* + *Hpa*) hydrolysate reveals that *Hpa* and *Hae* sites are also clustered with each other in about 60 (*Hae*,*Hpa*) site clusters. These "restriction site clusters" have a G+C-content in the 45 to 62% range and are about 35 base-pairs long on the average; "G+C-rich clusters", which have a G+C-content of 60%, and contain no *Hae* or *Hpa* site, are possibly adjacent to the site clusters. In addition to the *Hae* and *Hpa* sites which are present in restriction site clusters, 40 "isolated" *Hpa* sites and very few "isolated" *Hae* sites also exist on the mitochondrial genome.

The (*Hae*,*Hpa*) site clusters have nucleotide sequences similar to those found in bacterial promotors and operators, in that they consist of hyphenated palindromes in which short G-C and A-T sequences alternate with each other. Their localization is at the border of subsequent gene–spacer units, and their number may well be equal to that of mitochondrial genes and spacers. It is, therefore, very likely that the (*Hae*,*Hpa*) site clusters (and possibly, the G+C-rich clusters) are regulatory elements and that the mitochondrial genome of yeast is made up of genetic units, each one of which is formed by a regulatory sequence, a gene and a spacer.

1. Introduction

Mitochondrial DNA from yeast is characterized by a very striking compositional heterogeneity as first shown (Bernardi *et al.*, 1970) by its differential melting curve which exhibits: (1) a component representing about half of the DNA and having a very low t_m, lying between those of poly(dA·T-dA·T) and poly(dA,dT); (2) a number of components melting over a very extended temperature range (20 deg. C).

Micrococcal nuclease digestion combined with gel filtration (Prunell & Bernardi, 1974) showed that 50% of mitochondrial DNA are formed by long stretches (M_w $1{\cdot}6\times10^5$) with a G+C content of less than 5%, essentially consisting of short sequences of alternating and non-alternating dA,dT (Bernardi & Timasheff, 1970; Ehrlich *et al.*, 1972). The other 50%, probably containing the genes and their regu-

† Paper V is the preceding paper.

‡ Present address: Department of Biological Chemistry, Harvard Medical School, Boston, Mass. 02115, U.S.A.

§ To whom all correspondence should be addressed.

latory elements, have an average G+C content of 32%; they are, however, very heterogeneous in base composition, the G+C level ranging from 22% in the ribosomal cistrons (Reijnders & Borst, 1972) and 28 to 35% in the tRNA genes (Schneller *et al.*, 1975) to 65% in double-stranded fragments (M_w = 40,000) which can be prepared by micrococcal nuclease. These represent 10% of mitochondrial DNA (Prunell & Bernardi, 1974).

The present investigations show that this latter material largely comprises about 60 clusters of C-C-G-G and G-G-C-C sequences together with G+C-rich clusters (G+C = 60%) not containing such sequences. The C-C-G-G and G-G-C-C clusters have sequence properties similar to those of prokaryotic promotors and operators, are localized at the border of subsequent gene–spacer units, appear to be stoichiometric with mitochondrial genes and spacers, and are therefore likely to correspond to regulatory sequences. The G+C-rich clusters, which are possibly adjacent to the site clusters, might fulfil a similar role.

The main results of this work were presented at the Xth Meeting of the Federation of European Biochemical Societies (Paris, July 1975), at the 10th International Bari Conference on the Genetic Function of Mitochondrial DNA (Bernardi *et al.*, 1976), and at the Interdisciplinary Conference on the Genetics and Biogenesis of Chloroplasts and Mitochondria (Bernardi, 1976).

2. Materials and Methods

Materials and Methods were described in the preceding paper (Prunell *et al.*, 1977) except for the following: (i) degradations of mitochondrial DNAs with *Hha*I† were performed as with *Hpa*II or *Hae*III (Prunell *et al.*, 1977). (ii) Base-composition determination and degradation of mitochondrial DNA with micrococcal nuclease were performed as described by Prunell & Bernardi (1974). (iii) Double degradations with *Hpa* and *Hae* were performed in 3 different ways: (1) both enzymes were added together to the incubation mixture. (2) One enzyme was added first and the other one after 16 h; the completeness of the first degradation was checked by running a DNA mixture incubated with the first enzyme only; the procedure was used with either *Hpa* or *Hae* as the first enzyme added. (3) Single *Hpa* or *Hae* digests were obtained; these were deproteinized by shaking with chloroform/isoamyl alcohol, centrifuged and dialyzed against incubation solvent; *Hpa* or *Hae* DNA digests were then incubated with *Hae* or *Hpa*, respectively. Identical electrophoresis patterns of the digest were obtained in all cases.

3. Results

The probability of finding in a statistical DNA, that is a DNA in which the four nucleotides are distributed at random, the sequences G-G-C-C, C-C-G-G and G-C-G-C, which are the restriction sites of *Hae*, *Hpa* and *Hha*I, respectively (K. Murray & A. Morrison, unpublished results; Garfin & Goodman, 1974; Roberts *et al.*, 1976) is the same, $p = (G+C/2)^4$. The statistically expected number of such sites in each genome unit of yeast mitochondrial DNA, having N_0 nucleotide pairs, is $S = N_0 (G+C/2)^4 = 5$. Each genome unit releases, however, about 10 *Hha*I fragments (this work), and 71 to 113 *Hae* or *Hpa* fragments (Prunell *et al.*, 1977). These results indicate a highly non-random distribution of *Hae* and *Hpa* sites in yeast mitochondrial DNA. It is unlikely that these sites are present in the A+T-rich spacers ($S = 0{\cdot}015$) or that mitochondrial gene sequences deviate so much from

† Restriction endonucleases are indicated here according to Smith & Nathans (1973). *Hpa*II and *Hae*III will be indicated as *Hpa* and *Hae* in this paper. *Hha*I is from *Haemophilus haemolyticus*.

random. It is therefore possible that *Hae* and *Hpa* sites are clustered in special regions of the mitochondrial genome. Such a possibility has been explored by using two experimental approaches, namely hydroxyapatite chromatography of *Hae* and *Hpa* digests and gel electrophoresis of *Hae* + *Hpa* double digests. The first approach allows the detection of single-stranded oligonucleotides released because the restriction sites are so close together that the intervening duplex becomes unstable, the second provides the number of (*Hae*,*Hpa*) clusters.

(a) *Hydroxyapatite chromatography of DNA degraded by* Hae *or* Hpa

(i) The chromatogram of the *Hpa* digest of DNA from *Saccharomyces cerevisiae* (strain B) on hydroxyapatite is characterized (Fig. 1(a)) by the presence of material eluting at 0·05 M, 0·1 M and 0·15 M-sodium phosphate and by a broad peak eluting at 0·22 to 0·32 M-phosphate.

The main peak material was formed by double-stranded fragments as indicated by its elution molarity range, 0·22 to 0·32 M-phosphate (Bernardi, 1965*a*,1971; Wilson & Thomas, 1973) and ranged in G+C-content from 29% for the first-eluted fractions to 12% for the last-eluted ones.

The material eluting at low salt had the following properties:

(1) it was formed by single-stranded oligonucleotides, as indicated by its elution molarity range, 0·05 to 0·15 M-phosphate (Bernardi, 1965*a*,1971; Wilson & Thomas, 1973). Mononucleotides, which might arise from contamination of the restriction enzyme with exonuclease, would not be retained by a column equilibrated with 0·001 M-phosphate (Bernardi, 1964,1971), and were absent from the digest (see legend to Fig. 1(a));

(2) this material would not be seen in the gel electrophoresis experiments because of the poor binding of ethidium bromide by single-stranded oligonucleotides at the ionic strength used, and the lack of fluorescence of the dye bound to such structures (Le Pecq, 1971);

(3) it represented about 2% of the digest, assuming a hyperchromicity of 30%: this corresponds to 1×10^6 of DNA per 50×10^6 (M_r) genome unit;

(4) it was eluted in decreasing amounts by increasing molarity steps, showing its chromatographic discontinuity with the main peak material;

(5) it had a G+C content of 65%, showing its compositional discontinuity with the main peak material, which had a much lower G+C content (see above);

(6) it was absent in mitochondrial DNA preparations incubated in the absence of enzyme.

(ii) The chromatogram of the *Hpa* digest of DNA from *S. carlsbergensis* (strain C) on hydroxyapatite (Fig. 1(b)) was very similar to that just described for the DNA from strain B. The amount and base composition of the oligonucleotides released by *Hpa* were practically identical to those found for the oligonucleotides from the DNA of strain B (Table 1).

(iii) The chromatogram of a *Hae* digest from strain C on hydroxyapatite was basically similar to that just reported for *Hpa* digests. The single-stranded oligonucleotides eluted by 0·15 M-sodium phosphate (0·05 and 0·1 M steps were not used): (1) represented 1·7%, i.e. an M_r of $0{\cdot}8\times10^6$ of DNA per genome unit; (2) had a G+C content of 33% (Table 1); and (3) contained no mononucleotides. Very similar

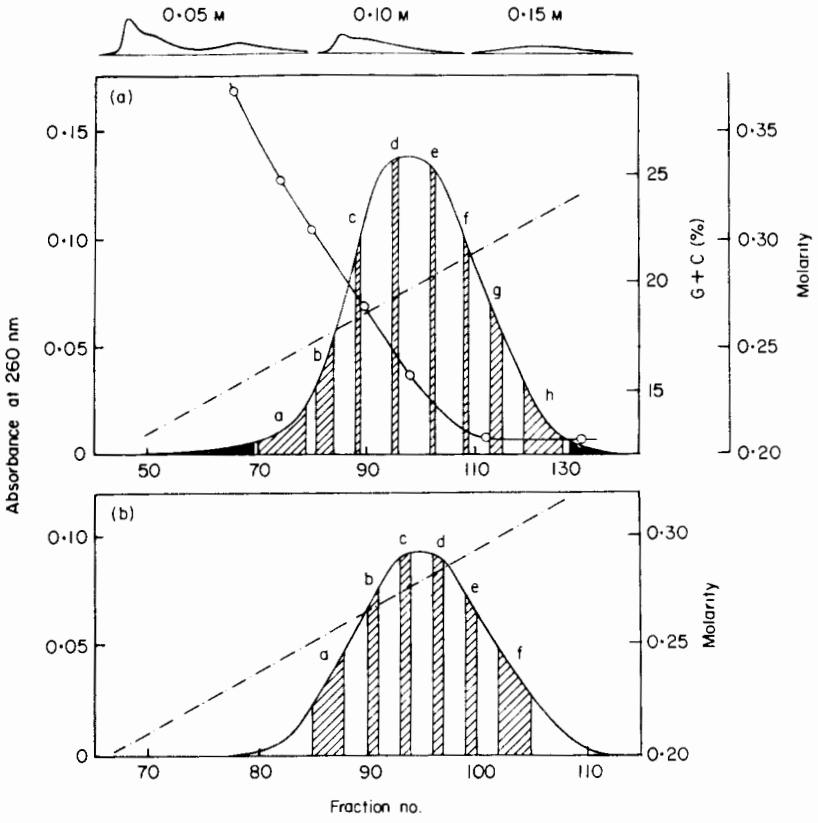

FIG. 1 (a) Chromatogram of 8·4 A_{260} units, in 6 ml TMSH (6·6 mM each of Tris (pH 7·4), $MgCl_2$, 2-mercaptoethanol, 10 mM-NaCl), of a *Hpa* digest of mitochondrial DNA from strain B, on a 0·95 cm × 22 cm hydroxyapatite column. Flow-rate was 16 ml/h; 2·1-ml fractions were collected. After degradation, no EDTA was added, and the digest was shaken with chloroform/isoamyl alcohol and centrifuged. After loading, the column was washed with the equilibration buffer, 0·001 M-sodium phosphate (pH 6·8), until the 2-mercaptoethanol of the DNA solvent was completely eluted. The corresponding peak (not shown in the Figure) was free from mono- and oligonucleotides, as demonstrated by chromatography on DEAE-cellulose. Steps of increasing molarities were then applied to the column. The 3 peaks above the frame are the enlarged transmission recordings, at 253·7 nm, of the material eluted at 0·05 M, 0·10 M and 0·15 M-phosphate, respectively. The corresponding amounts are 0·074, 0·060 and 0·050 A_{260} units. A linear gradient (0·15 to 0·40 M) of phosphate was then applied. (– · – · –) The phosphate molarities of the fractions. (○) The base composition. This was determined on pooled fractions in some cases (black areas). Hatched areas indicate fractions pooled and run on the 6% and 2·5% polyacrylamide gels (Figs 3 and 5).

(b) Chromatogram of 1·5 A_{260} units, in 2 ml TMSH, of a *Hpa* digest of mitochondrial DNA from strain C, on a 0·5 cm × 22 cm hydroxyapatite column. Flow-rate was 10 ml/h; 1·1-ml fractions were collected. Before loading, the sample was treated as described for (a). A linear gradient (0·1 to 0·4 M) of sodium phosphate (pH 6·8), was applied. The broken lines indicate the molarity. Hatched areas indicate the fractions pooled and run on the 2·5% polyacrylamide gel (Fig. 7). The material eluted below 0·2 M-phosphate is not shown.

TABLE 1

Amount and base composition of oligonucleotides released by Hae *and* Hpa *from the mitochondrial DNA of* S. carlsbergensis†

First degradation	Second degradation	Amount (%)	Amount (per genome unit $\times 10^{-6}$)	G + C (%)
Hpa	—	2·1	1·05	68‡
Hpa	*Hae*	0·5	0·25	37
Hae	—	1·7	0·85	33
Hae	*Hpa*	1·3	0·65	60
Hpa + *Hae*		2·6	1·3	62
Hae + *Hpa*		3·0	1·5	45

† Four A_{260} units of DNA were used in the first degradation: 3·5 A_{260} units in the second degradation. The amount of oligonucleotides was calculated from the A_{260}, assuming a 30% hyperchromicity.

‡ The oligonucleotides released by *Hpa* from the DNA of strain B represented 2% of the digest and had a G + C content of 65%.

results were obtained with a *Hae* digest from strain B; in addition, the base composition of the fractions from the main peak was determined and found to follow a pattern similar to that of *Hpa* digests (G. Martini, unpublished observations).

(b) *Hydroxyapatite chromatography of DNA sequentially degraded by* Hae *and* Hpa

Hpa or *Hae* digests, freed from the oligonucleotides released by these restriction enzymes by chromatography on hydroxyapatite, were submitted to a second digestion by *Hae* or *Hpa*, respectively. These digests were then run on hydroxyapatite; the oligonucleotides released by the second digestion were isolated; their amounts and base composition are presented in Table 1.

(c) Hpa *degradation of the G + C-rich fragments released from mitochondrial DNA by micrococcal nuclease*

Degradation with micrococcal nuclease (up to 90% degradation) was done at the melting temperature (Prunell & Bernardi, 1974). Separation of large fragments from short oligonucleotides was not done by gel filtration on Sephadex G100 columns, as in previous work, but by chromatography on hydroxyapatite. In this case (Fig. 2(a)) the fractions eluted at the steps 0·001 M, 0·05 M and 0·15 M-sodium phosphate correspond to the material which appeared in the retarded fraction of the G100 chromatogram (Prunell & Bernardi, 1974). The fraction eluted at the 0·4 M step, corresponding to the excluded fraction of the G100 experiment, was 9% of the total (on a nucleotide basis), and had a G + C content of 62%, in good agreement with previous data on DNA from strain A (Prunell & Bernardi, 1974).

The 0·4 M DNA fraction was then degraded with *Hpa* in standard conditions and re-chromatographed on hydroxyapatite (Fig. 2(b)). 15% (on a nucleotide basis) of this material was eluted by a 0·15 M step, the rest by a 0·4 M step. The G + C contents of these fractions were 74% and 60%, respectively.

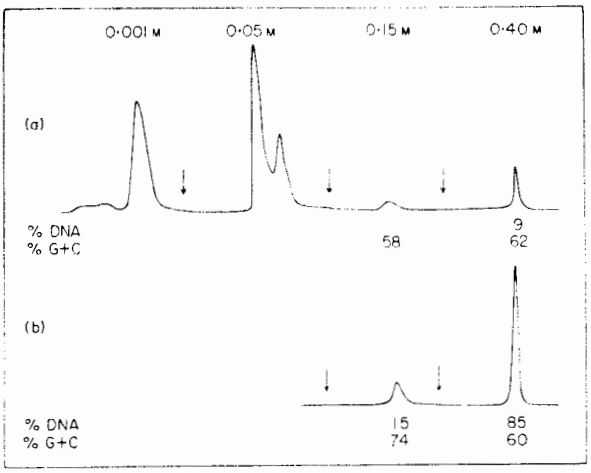

FIG. 2. (a) Hydroxyapatite chromatogram on a 0·84 cm × 8·5 cm column of 6 A_{260} units of mitochondrial DNA from strain B degraded by micrococcal nuclease at a temperature at which the hyperchromic shift was 51%. The column was previously equilibrated with 10^{-3} M-sodium phosphate buffer.

(b) Hydroxyapatite chromatogram, on a 0·84 cm × 3·9 cm column, of 0·250 A_{260} units of DNA recovered from the peak eluted at 0·4 M in the chromatogram from (a), and degraded with *Hpa*. The DNA was dialyzed against TMSH and degraded with *Hpa* in standard conditions. The 2-mercaptoethanol peak was not indicated.

In both chromatograms the arrows indicate the beginning of the steps; the columns were equilibrated with 10^{-3} M-sodium phosphate buffer (pH 6·8); the Figure presents transmission recordings at 253·7 nm.

The fact that the 0·15 M material obtained in this experiment had a G + C-content of 74%, whereas this was only 58% in the first one, rules out the possibility that the 0·15 M material of the second chromatogram was formed by denatured fragments incompletely removed by the first column.

(d) *Gel electrophoresis of* Hpa *digests previously fractionated on hydroxyapatite*

DNA from strains B and C was digested with *Hpa* and then fractionated on hydroxyapatite. Different fractions from the hydroxyapatite chromatogram were then separated into fragments by gel electrophoresis. The patterns show a number of interesting features (Figs 3 and 4): (1) all the small *Hpa* fragments are found in the fractions eluting first at low phosphate molarities. Among these the smallest show the normal elution behavior from hydroxyapatite in that their elution molarity increases with size. (2) In contrast some of the larger fragments show anomalous behavior. For example, fragments 16, 17 and 19 elute after 15; 12, 13, 14 after 11; 10 after 9 and so on. (3) Among the larger fragments there are other anomalies: the largest fragments do not elute at the highest molarity (lanes d, e, strain B; lanes a to d, strain C, Fig. 5). (4) There is also a striking similarity in the behavior of pairs of

Before electrophoresis, the hydroxyapatite fractions were pooled, concentrated to about 1 ml, dialyzed against 0·15 M-sodium chloride, 0·02 M-Tris·HCl (pH 7·4), adjusted to 1 M-sodium chloride, and precipitated with 3 vol. ethyl alcohol. The mixture was kept for at least 5 h at −20 °C and centrifuged at 0 °C for 1 h at 38,000 revs/min in a Spinco SW50.1 rotor. The supernatant was discarded and the pellets dried under reduced pressure. The DNA was recovered with 50 µl of electrophoresis buffer: 25 µl of each sample was loaded on the gel after being made 1% in sodium dodecyl sulfate and incubated at 37 °C for 30 min.

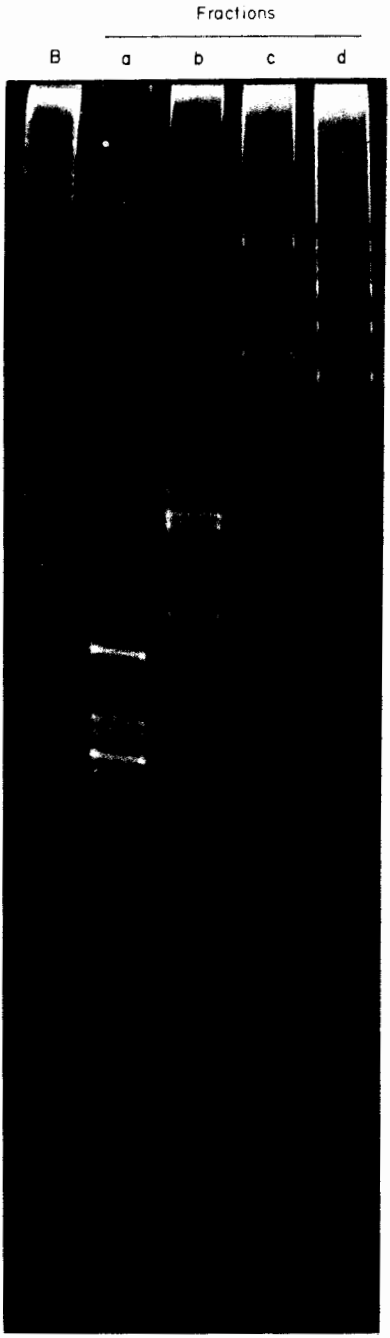

FIG. 3. Electrophoretic pattern on a 6% polyacrylamide gel of fractions eluting from the hydroxyapatite chromatogram of Fig. 1(a) at low phosphate molarities; letters refer to fractions indicated in Fig. 1(a). B corresponds to unfractionated DNA from strain B, degraded by *Hpa*. Slot a contained, besides fraction a, half of the material eluting between fractions 50 and 70 in Fig. 1(a).

(continued on facing page

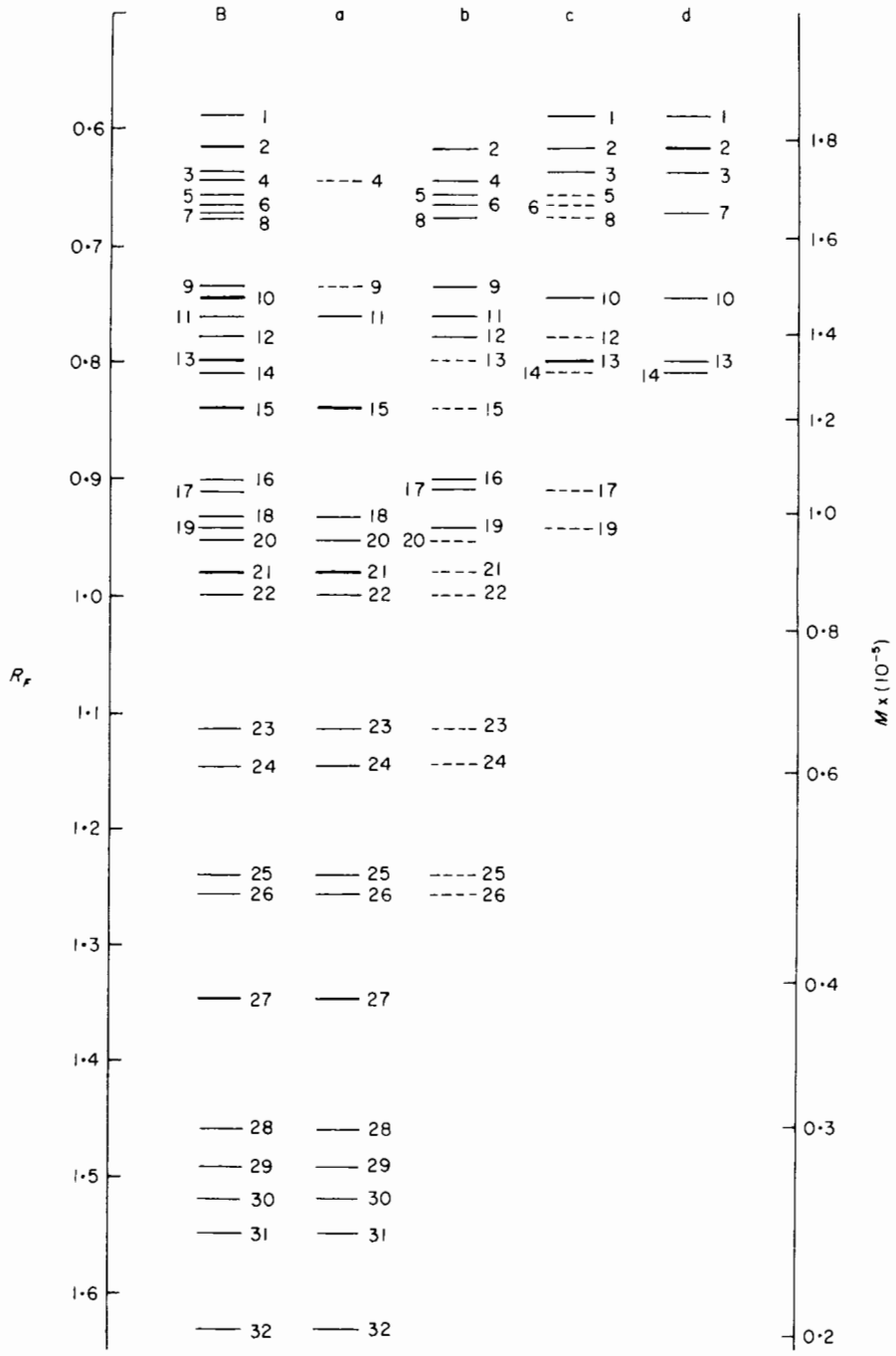

FIG. 4. Scheme of the band patterns of Fig. 3. Indications have the same meaning as in Fig. 3. The upper bands visible on Fig. 3 were not shown. It should be recalled (see legend to Fig. 3) that slot a contained, besides fraction a, half of the material eluting between fractions 50 and 70 in Fig. 1(a). Bands containing more than 1 fragment were indicated with thicker lines: weak bands were indicated by broken lines. Correspondence with bands of Fig. 12 of Prunell *et al.* (1977) are the following: bands 1 + 2 correspond to b26; 7 + 8 to B27; one fragment of 13 + 14 to B28; 17 + 18 to B29; and 21 + 22 to B30.

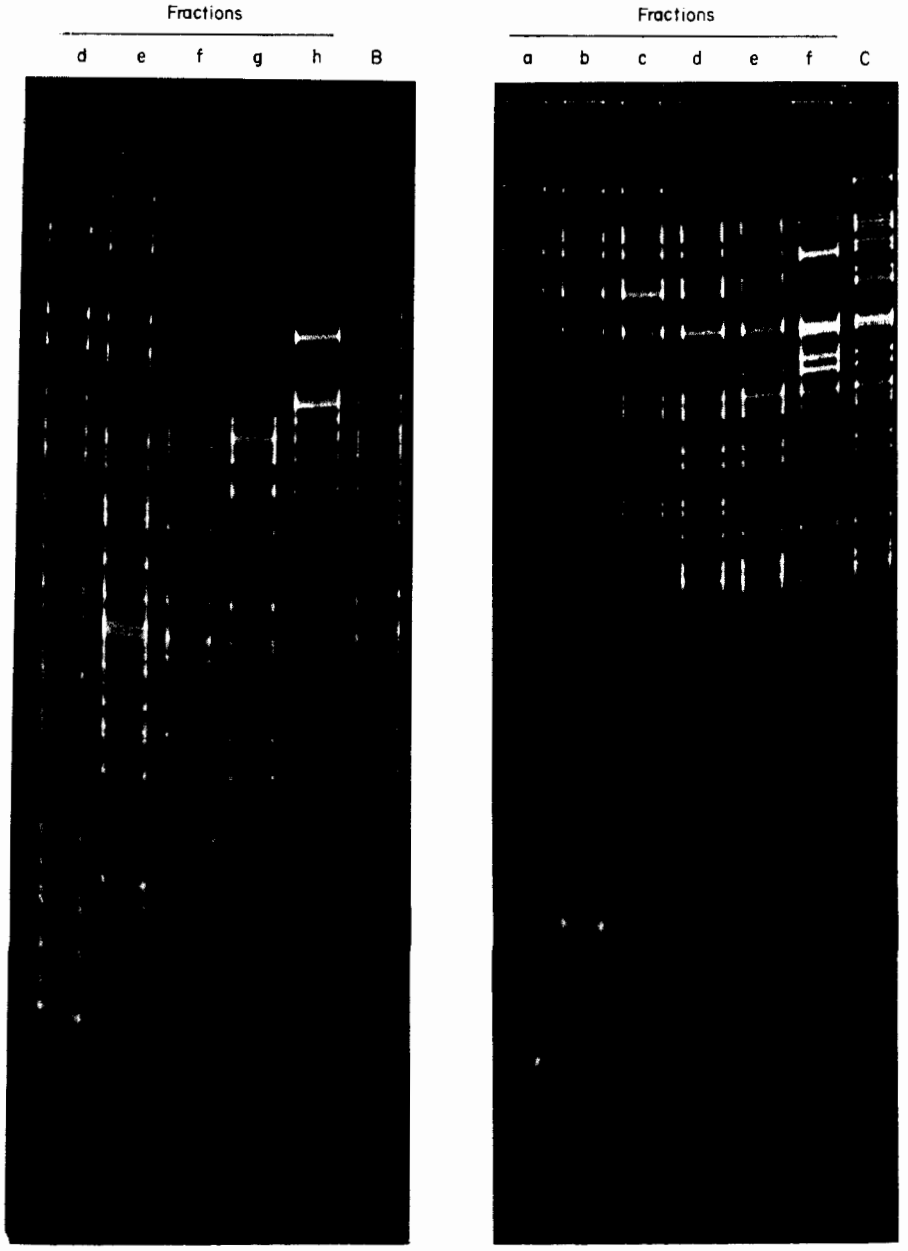

FIG. 5. Electrophoresis pattern on a 2·5% polyacrylamide/0·5% agarose gel of fractions from the hydroxyapatite chromatograms of Fig. 1(a),a and 1(b),b; the letters refer to the same components as in Fig. 3.

fragments between strains B and C. Two pairs, not the largest, $b6_2$, $b8_2$ (Fig. 6; these are the 7th and 11th fragments in DNA from strain B; lane h, Fig. 5(a)) and $c4_2$, $c8_2$ (Fig. 6; these are the 6th and 11th fragments in DNA from strain C: lane f, Fig. 5(b)) are characteristically eluted sharply at the end of the chromatogram; the

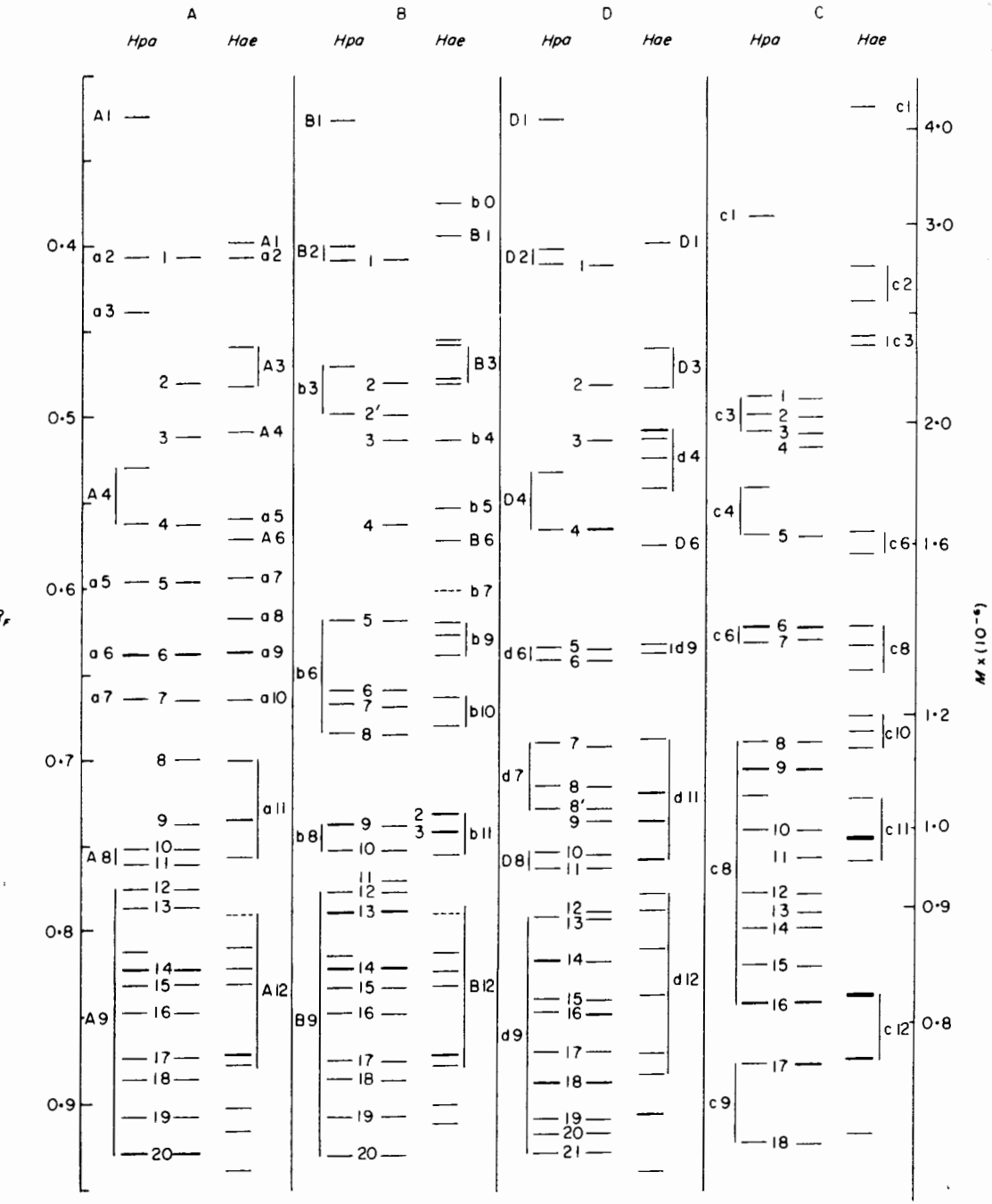

FIG. 6. Band patterns as obtained on the 2% gels with the four DNAs degraded by *Hpa* and *Hae* (simple digests) and *Hpa* + *Hae* (double digests). The bands or group of bands of simple digests are indicated as in Fig. 8 of Prunell *et al.* (1977). The bands of the double digests are numbered for convenience.

molecular weight differences between these two pairs of fragments are 1×10^7 and 6×10^5. The other nine highest molecular weight fragments from both DNAs are eluted as rather broad peaks at a relatively low molarity.

(e) *Analysis of* Hae. Hpa *double digests of mitochondrial DNAs*

Hae + *Hpa* double digests of the DNAs from the four yeast strains studied in the preceding paper (Prunell *et al.*, 1977) were studied on 0·5% agarose/2% polyacrylamide gels (Fig. 6). Since all double digests exhibited a very similar behavior on this gel, a complete analysis on gels of different concentration was only done for the *S. carlsbergensis* DNA (Figs 7 and 8).

(i) *General features*

Hae digestion of *Hpa* hydrolysates causes very few changes in the *Hpa* band patterns (Figs 6 and 8) since only a few *Hpa* bands (A1, B1, D1; $A4_1$, $D4_1$; one of the b8 fragments; $A9_3$, $B9_3$) are shifted down (Fig. 6); moreover, no additional fragments observable by gel electrophoresis are formed (Fig. 8). In sharp contrast, *Hpa* digestion of the *Hae* hydrolysate causes (1) an extensive change of the *Hae* band pattern which becomes very similar to the *Hpa* band pattern; in general, bands seem to move to positions corresponding to only slightly smaller molecular weights; (2) the appearance of a number of additional bands, most of which are in the lowest molecular weight range. Table 2 summarizes the findings obtained with the double digest of *S. carlsbergensis* DNA. About 80, 6 and 15% of double-digest fragments correspond, in mobility, to *Hpa* fragments, *Hae* fragments or to none of the single-digest fragments, respectively; the number of these three sorts of fragments, called in Table 2 "*Hpa*

TABLE 2

Results of double digestion (Hpa + Hae) *of mitochondrial DNA from* S. carlsbergensis

Single digests	*Hpa* fragments	107
	Hae fragments	71
Double digest	Fragments	107
	"*Hpa* fragments"	83
	"*Hae* fragments"	7
	"New fragments"	17
	Hpa fragments degraded by *Hae*	24
	Hae fragments degraded by *Hpa*	64

fragments", "*Hae* fragments" and "new fragments", respectively, are the starting point for the quantitative estimation of the different restriction sites and site clusters in mitochondrial DNA (see Appendix).

(ii) *Relationships between* Hpa *and double-digest bands*

A careful comparison of single and double digests (not presented here) allows the origin of double-digest bands to be traced back to the *Hpa* bands at least on the 2% gel (Fig. 6). A similar comparison cannot be done for the *Hae* digests because of the extensive mobility shifts occurring upon *Hpa* digestion. An exception is the

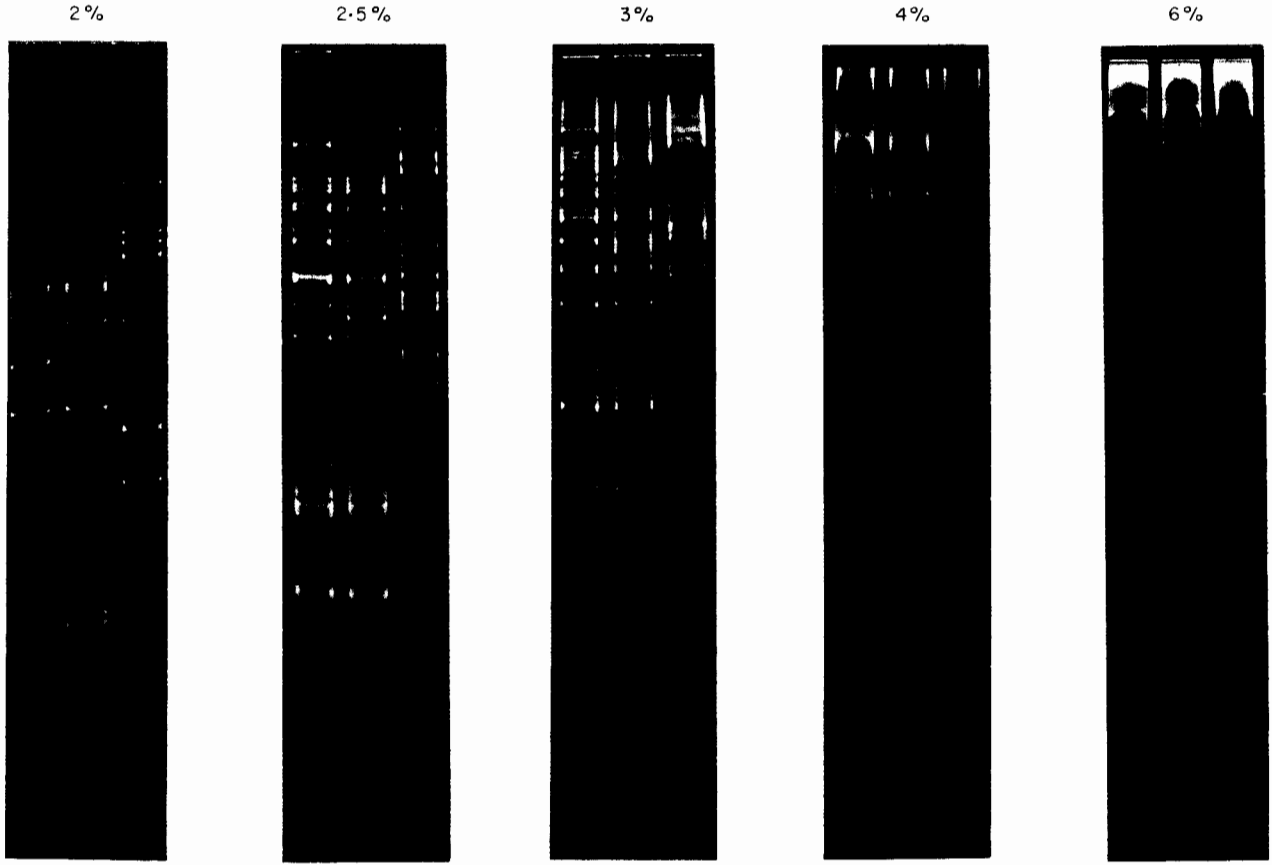

FIG. 7. Electrophoretic patterns of single and double digests of *S. carlsbergensis* DNA. Polyacrylamide concentrations are indicated. 0·5% agarose was present in all gels except for the 6% gel. In each case, *Hpa*, *Hae* + *Hpa*, and *Hae* digests are shown from left to right.

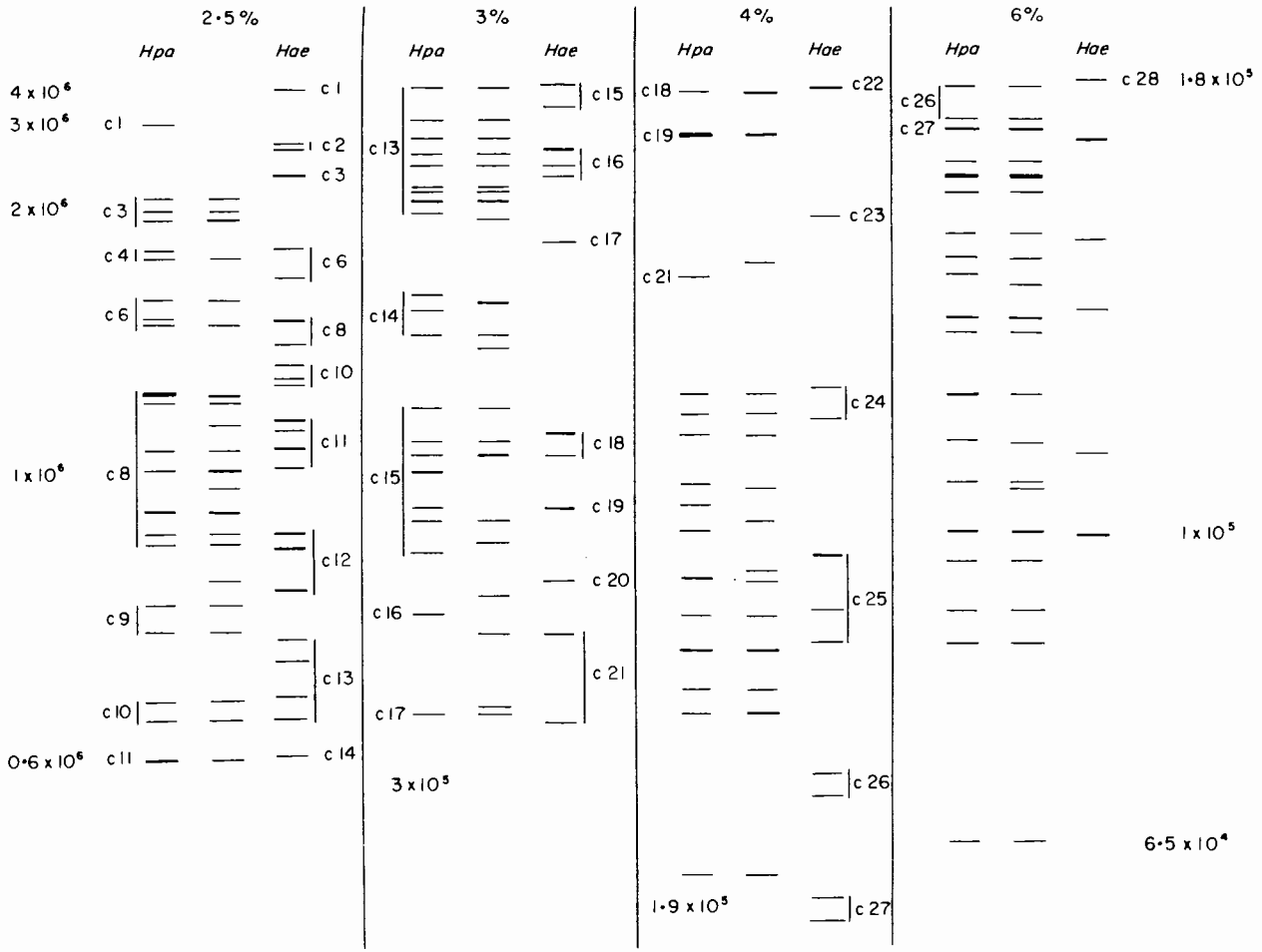

FIG. 8. Band patterns obtained on 2·5 to 6% polyacrylamide gels with *S. carlsbergensis* DNA degraded by *Hpa* and *Hae* (single digests) and *Hpa* + *Hae* (double digests). The bands or groups of bands of single digests are indicated as in Figs 9 to 12 of Prunell *et al.* (1977). Overlaps from gel to gel are not shown. Ten very low molecular weight bands having the same mobility in *Hpa* and *Hpa* + *Hae* digests, as well as 2 *Hae* bands, are not shown in the 6% gel.

case of fragments A1, B1 and D1 (*Hae*) which give rise, upon *Hpa* digestion, to the fragments labeled 1 in the A, B and D patterns (Fig. 6). This is an example of the shifts mentioned above (and corresponding in this case to a difference in molecular weight of $1{\cdot}3 \times 10^5$) which apparently occur in most *Hae* fragments upon *Hpa* degradation.

(iii) *Band homology*

Several *Hpa* bands from DNAs of different strains not only have the same mobility, but originate upon *Hae* digestion fragments having the same mobility, thus providing strong evidence for band homology. Such is the case for bands A1, B1, D1 which give rise to bands 2 + 3 in the double digest, and for bands $A4_1$ and $D4_1$ which originate bands 9.

4. Discussion

(a) *The* Hae *and* Hpa *restriction fragments*

Elution of yeast mitochondrial DNA fragments from hydroxyapatite depends (Piperno *et al.*, 1972) upon both molecular weight (smaller fragments being eluted first) and the presence of A+T-rich spacer sequences (these shift elution molarity to higher values). Under these circumstances, useful information on the *Hae* and *Hpa* fragments and their interstrain homology can be obtained from hydroxyapatite chromatography.

(i) Most of the small *Hpa* fragments ($M_r < 2 \times 10^5$) do not seem to contain spacer sequences, since they have a very high G+C-content (25 to 30%; compare Fig. 1(a) and Fig. 4), and are eluted from hydroxyapatite early and in an order corresponding to their molecular weight (Fig. 4); this is not the case for some other fragments which are shifted to higher elution molarities, probably because they do contain spacer sequences (see Results, section (d), (3)). The small *Hpa* fragments are present in about the same number in mitochondrial DNAs from different strains, they are the few fragments which are homologous between *S. cerevisiae* and *S. carlsbergensis* (Prunell *et al.*, 1977), and they do not correspond to any fragment seen in *Hae* digests. All the above findings suggest that the small *Hpa* fragments arise from isolated *Hpa* breaks in the mitochondrial genes (see also section (b) (v), below).

(ii) The bulk of the restriction fragments have a G+C content between 20 and 12% (Fig. 1(a)). Since "spacers" have a G+C-content lower than 5%, and "genes" an average G+C-content of about 26% (see section (d)), this finding suggests that most of the *Hae* or *Hpa* fragments contain roughly equal amounts of spacer and gene sequences.

(iii) The base composition of *Hpa* fragments eluting at increasing phosphate molarity is practically identical for *S. cerevisiae* and *S. carlsbergensis*; more significantly, gel electrophoresis of corresponding chromatographic fractions shows a very great similarity. Since elution molarity from hydroxyapatite depends upon the nucleotide sequence of DNA fragments (Piperno *et al.*, 1972), these data suggest a nucleotide homology between fragments having similar molecular weight. Interestingly in the case of the two pairs of fragments $b6_2$, $b8_3$ and $c4_2$, $c8_3$, the characteristic sharp elution pattern suggests a sequence homology and yet the molecular weights differ by 1×10^5 to 6×10^5; this indicates, in agreement with the conclusions of Prunell *et al.* (1977), that sequence homology exists between fragments of different length.

(b) *Restriction site clusters*

(i) *Hae* sites are clustered with each other; so are *Hpa* sites. Direct evidence for such clustering is provided by the finding in the DNA of *S. carlsbergensis* that *Hae* hydrolysates contain 1·7% and *Hpa* hydrolysates 2·1% single-stranded oligonucleotides (Table 1). These can only derive from the melting of DNA stretches which separate restriction sites and are too short to maintain a double-stranded configuration in our experimental conditions.

The possibility that the single-stranded oligonucleotides are originated by breaks near the ends of the starting mitochondrial DNA fragments or by a DNase contaminating the restriction enzyme used is ruled out by the G + C content of the oligonucleotides released by *Hpa* or *Hae*, which is very far from 18%, the G + C content of mitochondrial DNA.

(ii) A clustering of *Hae* and *Hpa* sites with each other, first suggested by the similar size distributions of the fragments produced by *Hae* or *Hpa* from a given mitochondrial DNA (Prunell *et al.*, 1977; Fig. 13), is shown by two main lines of evidence: (1) the *Hae* + *Hpa* double-digest of *S. carlsbergensis* DNA exhibits 107 fragments on the gel, whereas the total number of *Hae* + *Hpa* sites is at least 178 (71 + 107): *Hae* digestion of the *Hpa* hydrolysate does not cause the formation of additional fragments observable by gel electrophoresis, but releases single-stranded oligonucleotides, indicating that all *Hae* sites (with a few exceptions) are clustered with *Hpa* sites; the number of *Hae* fragments being 71, we can conclude that the number of (*Hae*,*Hpa*) site clusters is close to 60; (2) the results of Table 1 indicate that the amount of oligonucleotides released by *Hae* or *Hpa* in the second digestion is much lower than that released in the first one: *Hae* releases 0·5% of DNA after this has undergone *Hpa* degradation *versus* 1·7% released upon a first digestion; *Hpa* releases 1·3% of DNA after this has undergone *Hae* degradation *versus* 2·1% released upon a first digestion; this indicates that *Hae* clusters are contained, to a large extent, in *Hpa* clusters and *vice versa*.

(iii) The fact that the double-digest data indicate that essentially all *Hae* sites are clustered with *Hpa* sites, and the finding that 0·5% of DNA can be released by *Hae* from a *Hpa* hydrolysate, taken together, point to the existence of two classes of (*Hae*,*Hpa*) site clusters; the first of these (*Hpa-Hae-Hpa*) has *Hae* sites inside *Hpa* sites, whereas the second (*Hae-Hpa*) has *Hae* sites next to *Hpa* sites. The numerical estimation of each class of sites, given in the Appendix, indicates that the (*Hae*, *Hpa*) site clusters are almost equally distributed in the two sorts of clusters just mentioned.

(iv) Since the *Hpa* sites not clustered with *Hae* sites (see next subsection) are not likely to release any oligonucleotides, the data of Table 1 indicate that 2·6 to 3% of mitochondrial DNA, namely with 1·3 to $1{\cdot}5 \times 10^6$ molecular weight per genome unit, is present in (*Hae*,*Hpa*) site clusters. Since the number of the latter is about 60, each site cluster is formed, on the average, by 35 base-pairs. Now, we know that the smallest native fragments separated by gel electrophoresis are only 15 base-pairs long (Prunell *et al.*, 1976) and they are stable in spite of their G + C content of 29% (see Results, section (a)). Because of the higher thermal stability of restriction site clusters (G + C = 45 to 62%), it is unlikely that these melt unless they become shorter than about 10 base-pairs, which implies that each site cluster contains on the average no less than 4 to 5 sites (see Fig. A2 of Appendix), corresponding to 16 to 20 base-pairs. An implication of this conclusion is that the number of restriction sites for

both *Hpa* and *Hae* is much larger than that of restriction fragments. A very conservative estimate may be about 200 *Hpa* sites and about 150 *Hae* sites; both values are very much higher than expected on a statistical basis (see Results). If one takes into account the fact that the sites have a G+C level of 100%, and that the clusters have a level of 45 to 62%, this means that most of the base-pairs alternating with the restriction sites in the clusters must be A·T base-pairs. In this connection, it should be mentioned that the number of *Sma*I sites (C-C-C-G-G-G; C. Mulder, personal communication) on yeast mitochondrial DNA is very small if not zero (H. Kopecka, unpublished observations). This suggests that the nucleotides flanking the very many *Hpa* sites, C-C-G-G, are not random ones.

(v) Indications exist (see section (a) (i)) that the *Hpa* sites which are not clustered with *Hae* sites, that is isolated *Hpa* sites, are located within mitochondrial gene sequences. If such is the case, it is very unlikely that they are present in *Hpa* clusters (see Appendix). An attractive possibility for the isolated *Hpa* sites is that they are present in the tRNA genes and correspond to self-complementary sequences in the stems of these RNAs. Finally, as already mentioned, a very small number of isolated *Hae* sites also exist.

(c) *The G+C-rich clusters*

The proportion of restriction site clusters (about 3% of DNA) indicates that these only account for a fraction of the G+C-rich fragments released by micrococcal nuclease (Prunell & Bernardi, 1974), which form about 9% of the DNA. The rest constitutes what we will call here the "G+C-rich clusters". These are resistant to *Hpa* and *Hae*, otherwise small fragments derived from them would have been detected; they must, therefore, be present in the restriction fragments seen after gel electrophoresis. They are also resistant to *Hha*I, because only a very small number of fragments (~10) are released by this enzyme. Finally, the G+C-rich clusters are probably contiguous to the (*Hae*,*Hpa*) site clusters; in fact, *Hpa* degradation of the G+C-rich fragments released by micrococcal nuclease liberated single-stranded oligonucleotides in a yield (1·4%; i.e. 15% of the material representing 9% of total DNA; Fig. 2) rather close to that obtained from total DNA (2·1%).

(d) *The significance of restriction site clusters and of G+C-rich clusters*

(i) The (*Hae*,*Hpa*) restriction site clusters are characterized by (1) a high concentration in two symmetrical nucleotide sequences (G-G-C-C and C-C-G-G) which are complementary palindromes; and (2) the alternation of short A-T and G-C sequences. The symmetrical nucleotide sequences suggest by themselves a binding role for proteins having a dyad axis of symmetry. Such a symmetry relationship (Bernardi, 1965*b*,1968) has been found in all specific protein–DNA interactions (see Sobell, 1973, for a review). Both features found in the (*Hae*,*Hpa*) site clusters are characteristically present in all promotor and operator sites of phage and bacterial DNAs sequenced so far (see Dykes *et al.*, 1975, for a review). Sequencing will tell whether the *Hae* and *Hpa* sites are symmetrically arranged in the clusters.

The number of (*Hae*,*Hpa*) restriction site clusters, about 60, in the *S. carlsbergensis* mitochondrial DNA, is compatible with the coding capacity of the mitochondrial genome of yeast, since it corresponds to average gene sizes of about 350,000, which is not too low a figure for a genome in which the small tRNA genes represent such a large percentage (over 30%) of the total. The presently known mitochondrial genes

comprise 24 or more tRNA genes (Reijnders & Borst, 1972; Casey *et al.*, 1972), two rRNA genes (Reijnders *et al.*, 1972) and at least nine genes coding for polypeptide chains (Schatz & Mason, 1974; Schatz, 1976).

The (*Hae,Hpa*) clusters are localized at the border of subsequent gene–spacer units. In fact they cannot be localized (a) within the spacers, since in such a case they would be completely released by micrococcal nuclease at 50% digestion, a possibility contradicted by previous results (Prunell & Bernardi, 1974; Fig. 7); (b) within the genes, since each gene would then contain such peculiar sequences; (c) at the border of each gene and each spacer, since this would release material corresponding in composition to the spacers.

The nucleotide sequences, the number and the localization of (*Hae,Hpa*) site clusters strongly suggest a promotor and/or operator role; in this connection, it should be mentioned that (1) the absence of spacer transcripts in wild-type yeast mitochondria suggests a monocistronic transcription; the considerable distance between the sequences specifying the large and small ribosomal RNAs (Sanders *et al.*, 1975) also implies an independent transcription of these genes; (2) phenomena of genomic repression and induction are well known in yeast mitochondria; the existence of repressor protein(s) (of nuclear or mitochondrial origin) binding to operator sites to shut off transcription is a distinct possibility.

Another likely role of the (*Hae,Hpa*) site clusters is that they contain initiation sites for DNA replication; in fact, it is well known that the defective genomes of "petite" mutants can originate, by a deletion mechanism, from any region of the wild-type genome; this suggests that DNA replication can be initiated at multiple sites on the latter.

If such suggestions are correct, the expression and replication of the mitochondrial genome occur very differently in yeast and animal cells; in the latter case the two mitochondrial DNA strands are transcribed into RNAs (Aloni & Attardi, 1971,1972; Murphy *et al.*, 1975), which are then processed to yield rRNAs, tRNAs and mRNAs, and there is a single site for the initiation of DNA replication.

(ii) G+C-rich clusters. These are very probably formed from non-random nucleotide sequences, since they do not contain any of the quadruplets C-C-G-G, G-G-C-C or G-C-G-C, in spite of their high G+C level. It is likely that the G+C-rich clusters are responsible for the abundance of pyrimidine isostichs C_3, C_3T and C_4 in mitochondrial DNA (Ehrlich *et al.*, 1972). If most of the G+C-rich fragments released by micrococcal nuclease (which represent 9% of the DNA) are formed by restriction site clusters (about 3%) and G+C-rich clusters, the G+C-rich clusters contain more DNA than the (*Hae,Hpa*) clusters. As far as the localization of the G+C-rich clusters is concerned, the points raised for the (*Hae,Hpa*) clusters also apply, except that it is not ruled out that the G+C clusters exist on both sides of spacers; furthermore their distribution in length is unknown; they might therefore be contiguous to all or only to a fraction of restriction site clusters. Nothing can be said about the role of the G+C-rich clusters except that it is possible that they share the roles discussed above for the (*Hae,Hpa*) site clusters. It should be mentioned that, if one subtracts the contribution of (*Hae,Hpa*) site clusters and of G+C-rich clusters, the average G+C content of mitochondrial genes can be estimated to lie close to 26%.

(iii) If the stoichiometry and the topology of the genome elements is that suggested above, then the mitochondrial genome of yeast is organized in a number of genetic units, each one of which contains the four sequence elements depicted in

Figure 9 and has an average size of about 1200 base-pairs or a molecular weight of 8×10^5. It should be stressed that Figure 9 is only intended to give a general idea of the organization of the mitochondrial genome of yeast: in particular, the localization and the length of the G + C-rich clusters is still a matter of conjecture.

	G+C-rich cluster	Site cluster	Gene	Spacer
b.p. (av.)	~50	~35	~530	~600
G+C	60%	45 to 62%	26%	< 5%

FIG. 9. Hypothetical scheme of the organization of genetic units from the mitochondrial genome of yeast. The relative order of G + C-rich and restriction site clusters is unknown. The length of the G + C-rich clusters assumes a uniform distribution of these sequences near the restriction site clusters. b.p. (av.) Average size in base-pairs.

In spite of the uncertainty in a number of details, the picture of the organization of the yeast mitochondrial genome which is emerging from the present work is typically eukaryotic. The basic feature is an interspersion of unique sequences, devoted to the coding of rRNAs, tRNAs and mRNAs, with repetitive sequences, devoted to a regulatory and to a recombinational role: internal sequence repetitions in both spacers, restriction and G + C-clusters could allow unequal crossing-over events to take place in these sequence elements (Prunell *et al.*, 1976: manuscript in preparation).

APPENDIX

An Analysis of the *Hae*III, *Hpa*II Double Digest of Mitochondrial DNA from *S. carlsbergensis*

(i) Three preliminary remarks should be made: (1) whereas the patterns of single (*Hae* or *Hpa*) digests do not permit us to distinguish isolated (or non-clustered) from clustered breaks, the patterns of double digests permit us to do so by comparison with the simple digests: this is essentially due to the similarity of the double digest and of the *Hpa* digest; it should be emphasized that in the following analysis only single *Hae* or *Hpa* breaks will be counted, even if the effective cleavage of the DNA has been caused by several clustered breaks; (2) three classes of fragments are present in the double digest: (1) "*Hae* fragments" having the same mobility as fragments of the *Hae* digest; (2) "*Hpa* fragments" having the same mobility as fragments of the *Hpa* digest; (3) "new fragments" corresponding to none of the fragments in the single digests. Fragments having the same mobility in the *Hpa* and in the double digest certainly have the same nucleotide sequences, since only a few fragments of the *Hpa* digest change their mobility upon *Hae* digestion. The same cannot be said of "*Hae* fragments" which might, therefore, belong, in part at least, in the other two classes; the uncertainty about the number of "*Hae* fragments" is not important, however, for what follows, since their number is small. As far as argument (2) is concerned, it is important to know the nature of the ends of the

fragments of the double digest. By definition, each "new fragment" has a *Hae* end and a *Hpa* end; two *Hpa* and two *Hae* ends will be attributed to "*Hpa* fragments" and "*Hae* fragments", respectively. It should be stressed, however, that if a *Hae* site is extremely close to a *Hpa* site the fragment shift due to the second digestion may escape detection. Occurrence of such undetectable "new fragments" will cause a decrease in the estimated number of *Hpa-Hae-Hpa* clusters and an increase in the estimated number of *Hpa-Hae* clusters (see below), the overall number being, however, the same. (3) In the following analysis we will use a value of 71 as the number of *Hae* fragments and we will assume that all *Hae* sites are clustered with *Hpa* sites. In fact the number of *Hae* fragments is between 71 and 63 (Prunell *et al.*, 1976)

Hpa breaks in *Hae* fragments	1	2	3
"*Hpa* fragments" in the double digest	0 0	1 1 1	2 2 2
"New fragments" in the double digest	2 1	2 1 0	2 1 0
Type of fragment	(a) (b)	(c) (d) (e) (f)	

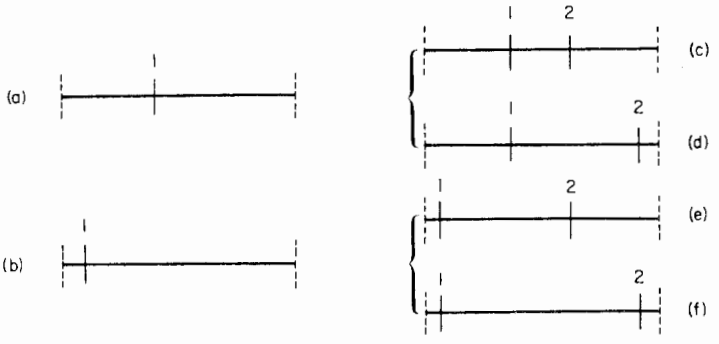

FIG. A1. Model of the degradation of the mitochondrial DNA *Hae* fragments by *Hpa* endonuclease. Broken and solid lines indicate *Hae* and *Hpa* breaks, respectively; 1 and 2, first and second cleavages.

and a few *Hae* sites are not clustered with *Hpa* sites (see Results, section (f), main text). This will, however, not change the numerical results very much.

(ii) As shown in Table 2, 64 out of 71 *Hae* fragments are degraded by *Hpa*, originating 17 "new fragments" and 83 "*Hpa* fragments". A single *Hpa* break on each *Hae* fragment originates two "new fragments" (Fig. A1(a)), one of which can be lost in the electrophoretic pattern if the *Hpa* break is close to one end (Fig. A1(b)). An additional *Hpa* break, not clustered with the previous one, generates a "*Hpa* fragment" (Fig. A1(c), (d), (e), (f)). This may lead to the elimination of one of the "new fragments" originated by the first break (Fig. A1(d), (f)).

The 17 "new fragments" can therefore derive from a number of *Hae* fragments comprised between 9 (Fig. A1(a) and (c)) and 17 (Fig. A1(b), (d), (e)). The remaining

55 (64 − 9) to 47 (64 − 17) fragments do not originate "new fragments" (Fig. A1(f)). More *Hpa* breaks may go into both sorts of fragments, generating "*Hpa* fragments".

In the first case (17 "new fragments" deriving from 9 *Hae* fragments) it can be seen that the 9 *Hae* fragments receive a minimum of 9 *Hpa* breaks not clustered with *Hae* breaks (Fig. 1(a)) and that the 55 (64 − 9) remaining fragments receive 110 *Hpa* breaks clustered with *Hae* breaks (Fig. 1(f)) and possibly some non-clustered *Hpa* breaks generating "*Hpa* fragments". These minimum estimates lead to 55 "*Hpa* fragments", whereas 83 have been observed. Therefore 28 (83 − 55) additional non-clustered *Hpa* breaks must occur, leading to a total number of 147 (9 + 110 + 28) *Hpa* breaks, 110 of which are clustered with *Hae* breaks.

In the second case (17 new fragments deriving from 17 *Hae* fragments) the same number of *Hpa* breaks (clustered and not clustered with *Hae* breaks, respectively) is found

Only 107 *Hpa* fragments out of 147 were seen in the single *Hpa* digest; 40 (147 − 107) fragments were therefore lost. This leads to 80 *Hpa* breaks clustered in pairs with each other, in 40 *Hpa* clusters. (That the number of the clustered *Hpa* breaks is 80 and not between 40 and 80 is evident if one thinks that the *Hpa* clusters are necessarily interspersed. Other *Hpa* clusters adjacent to each other could exist however; see below.) 110 *Hpa* sites are clustered with *Hae* sites; all the *Hpa* clusters contain therefore an internal *Hae* site. Of the remaining 67 (147 − 80) *Hpa* sites, 37 (147 − 110) are isolated and 30 (110 − 80) are clustered with *Hae* sites (*Hpa-Hae* clusters).

(iii) Some of the above figures can be checked as follows. Upon *Hae* digestion of *Hpa* fragments, only 24 fragments are degraded originating 7 "*Hae* fragments" and 17 "new fragments". By reasoning as above, it can be seen that 17 "new fragments" can only derive from 17 *Hpa* fragments, the 7 (24 − 17) "*Hae* fragments" deriving from the remaining 7 *Hpa* fragments. Only 31 (7 × 2 + 17) *Hae* breaks are, therefore, responsible for the degradation of the 24 *Hae* fragments; the corresponding 31 *Hae* sites are clustered with *Hpa* sites in agreement with the value given above. 40 (71 − 31) *Hae* breaks non-clustered with each other must therefore go into 40 small fragments released from the *Hpa* digest and become lost, again in agreement with the conclusion given above.

(iv) Figure A2(a), (b) and (c) summarizes the results obtained. As indicated above, this analysis does not allow us in general to distinguish one site from several clustered

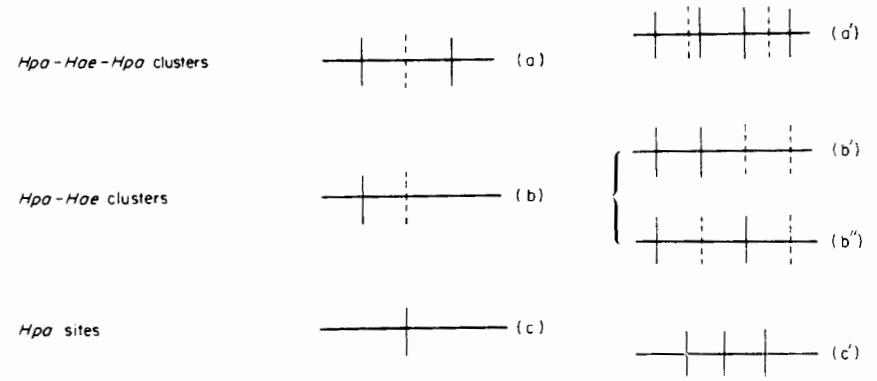

FIG. A2. Broken and solid vertical lines have the same meaning as in Fig. A1. The site number in (a'), (b'), (b'') and (c') is arbitrary.

ones (these sites being of the same nature) and therefore does not differentiate situation (a′) from (a), (b′) or (b″) from (b), and (c′) from (c). However, analysis of restriction digests on hydroxyapatite columns (section (d) in Results) not only confirms the existence of *Hpa* clusters (situations (a′) or (a), (b′) or (b″)) but shows that *Hae* clusters exist which are partly contained in *Hpa* clusters (Fig. A2(a′) and/or (b″)). On the other hand, considerations presented in the Discussion lead us to suppose that situation (c′) of Figure A2 is rare, and that the predominant one is situation (c).

We wish to thank our colleagues Helena Kopecka and Alberto Bernardi for the gift of *Hpa*II and *Hae*III, Godeleine Fonty for the gift of mitochondrial DNA preparations. Dr R. Roberts (Cold Spring Harbor Laboratories) kindly provided a sample of *Hha*I.

REFERENCES

Aloni, Y. & Attardi, G. (1971). *Proc. Nat. Acad. Sci., U.S.A.* **68**, 1757–1761.
Aloni, Y. & Attardi, G. (1972). *J. Mol. Biol.* **70**, 363–373.
Bernardi, G. (1964). *Biochim. Biophys. Acta*, **91**, 686–688.
Bernardi, G. (1965*a*). *Nature* (*London*), **206**, 779–783.
Bernardi, G. (1965*b*). *J. Mol. Biol.* **13**, 603–605.
Bernardi, G. (1968). *Advances in Enzymology* (Nord, F. F., ed.), vol. 31, pp. 1–49. Interscience, New York.
Bernardi, G. (1971). *Methods in Enzymology* (Grossman, L. & Moldave, K., eds), vol. 21, pp. 95–139. Academic Press, New York.
Bernardi, G. (1976). In *The Genetics and the Biogenesis of Chloroplasts and Mitochondria* (Bücher, T. *et al.*, eds), Elsevier-North-Holland, Amsterdam; in the press.
Bernardi, G. & Timasheff, S. N. (1970). *J. Mol. Biol.* **48**, 43–52.
Bernardi, G., Faurès, M., Piperno, G. & Slonimski, P. P. (1970). *J. Mol. Biol.* **48**, 23–42.
Bernardi, G., Piperno, G. & Fonty, G. (1972). *J. Mol. Biol.* **65**, 173–189.
Bernardi, G., Prunell, A. & Kopecka, H. (1975). In *Molecular Biology of Nucleocytoplasmic Relationships* (Puiseux-Dao, S., ed.), pp. 85–90. Elsevier, Amsterdam.
Bernardi, G., Prunell, A., Fonty, G., Kopecka, H. & Strauss, F. (1976). In *The Genetic Function of Mitochondrial DNA* (Saccone, C. & Kroon, A. M., eds), pp. 185–198. Elsevier-North-Holland, Amsterdam.
Casey, J., Cohen, M., Rabinowitz, M., Fukuhara, H. & Getz, G. S. (1972). *J. Mol. Biol.* **63**, 431.
Dykes, G., Bambara, R., Marians, K. & Wu, R. (1975). *Nucleic Acids Res.* **2**, 327–345.
Ehrlich, S. D., Torti, G. & Bernardi, G. (1971). *Biochemistry*, **10**, 2000–2009.
Ehrlich, S. D., Thiery, J. P. & Bernardi, G. (1972). *J. Mol. Biol.* **65**, 207–212.
Garfin, D. E. & Goodman, H. M. (1974). *Biochem. Biophys. Res. Commun.* **59**, 108–116.
Le Pecq, J. B. (1971). In *Methods of Biochemical Analysis* (Glick, D., ed.), vol. 20, p. 41–86. John Wiley, New York.
Murphy, W. I., Attardi, B., Tu, C. & Attardi, G. (1975). *J. Mol. Biol.* **99**, 809–814.
Piperno, G., Fonty, G. & Bernardi, G. (1972). *J. Mol. Biol.* **65**, 191–205.
Prunell, A. & Bernardi, G. (1974). *J. Mol. Biol.* **86**, 825–841.
Prunell, A., Kopecka, H., Strauss, F. & Bernardi, G. (1977). *J. Mol. Biol.* **110**, 17–52.
Reijnders, L. & Borst, P. (1972). *Biochem. Biophys. Res. Commun.* **47**, 126–133.
Reijnders, L., Kleisen, C. M., Grivell, L. A. & Borst, P. (1972). *Biochim. Biophys. Acta*, **272**, 396–407.
Roberts, R. J., Myers, P. A., Morrison, A. & Murray, K. (1976). *J. Mol. Biol.* **103**, 199–208.
Sanders, J. P. M., Heyting, C. & Borst, P. (1975). *Biochem. Biophys. Res. Commun.* **65**, 699–707.
Schatz, G. (1976). In *Mitochondrial Genetics, Biogenesis and Bioenergetics* (Bandlow, W. *et al.*, eds), pp. 163–177. De Gruyter, Berlin.
Schatz, G. & Mason, T. L. (1974). *Annu. Rev. Biochem.* **43**, 51–87.

Schneller, J. M., Faye, G., Kujawa, C. & Stahl, A. J. C. (1975). *Nucleic Acids Res.* **2**, 831–838.

Smith, H. & Nathans, D. (1973). *J. Mol. Biol.* **81**, 419–423.

Sobell, H. M. (1973). In *Advances in Genetics* (Caspari, E., ed.), vol. 17, pp. 411–490. Academic Press, Inc., New York.

Wilson, D. A. & Thomas, C. A., Jr (1973). *Biochim. Biophys. Acta*, **331**, 333–340.