Nucleotide sequence of a highly repetitive component of rat DNA

Michael Pech, Tibor Igo-Kemenes and Hans G. Zachau

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

Received 6 August 1979

ABSTRACT

A highly repetitive component of rat DNA which could not yet be enriched by density gradient centrifugation was isolated with the help of the restriction nuclease Sau3AI. This nuclease converted the bulk of the DNA to small fragments and left a repetitive DNA component as large fragments which were subsequently purified by gel filtration and electrophoresis. This DNA component which was termed rat satellite DNA I is composed of tandemly repeated 370 bp blocks. According to sequence analysis the 370 bp repeats consist of alternating 92 and 93 bp units with homologous but not identical sequences. Methylation of CpG residues was correlated to the rate of cleavage by restriction nucleases. Significant homologies exist between the sequences of rat satellite DNA I and satellite DNAs of several other organisms. The divergence of the sequence of rat satellite DNA I was discussed with respect to evolutionary considerations.

INTRODUCTION

Highly repeated DNA sequences have been found in the genomes of most, if not all eukaryotic organisms. Because of their ubiquity they have attracted considerable interest, although their functional significance is still unknown (reviews 1-4). The highly repeated DNA components of the rat genome have not been characterized in detail yet. On the basis of renaturation kinetics (e.g. 5) 8-10 % of the rat genome was determined to be highly repetitive. Conventional density gradient centrifugation, however, has failed to separate satellite DNA fractions, but separable components have been seen in alkaline CsCl gradients (6), in neutral CsCl gradients after DNA restriction (7) and in CsCl-netropsin density gradients (8). The analysis of rat DNA with restriction nucleases and gel electrophoresis has revealed a series of sharp bands indicating the presence of repetitive

DNA components (9,10). In previous experiments we mapped on a repetitive DNA component the cleavage sites for EcoRI and HindIII (11). These results were recently confirmed by Fuke and Busch (12) and extended to include the HaeIII and HinfI sites. Moreover in fractionation experiments these authors found most of the DNA component to occur in the nucleolar fraction (13).

In the course of our studies on the domain structure of chromatin in rat liver nuclei (14) we isolated a soluble chromatin fraction which contained this highly repetitive DNA in essentially pure form (11,15). In the present paper we report on the isolation of this DNA component by a different method and on its nucleotide sequence.

Although highly repetitive DNA components have usually been called satellite DNAs only when they could be separated by centrifugation techniques it seems justified to use for this thoroughly characterized DNA component which resembles closely other satellite DNAs the designation rat satellite DNA I. This term which has also the advantage of brevity will be used throughout the paper.

MATERIALS AND METHODS

Restriction nucleases. HindIII restriction nuclease was prepared according to (16 and H. Smith, pers. communication). Sau3AI (17) and HinfI were gifts of R.E. Streeck, Sau96I (18) was a gift of W. Hörz. HphI was purchased from New England Biolabs Inc., Beverly MA. Other enzymes and chemicals were as described previously (11). Digestions with restriction nucleases were carried out in buffer A (65 mM NaCl, 60 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 0.2 mM EDTA, 0.2 mM EGTA, 5 mM 2-mercaptoethanol, 15 mM Tris-HCl, pH 7.4) in the presence of 7 mM MgCl₂. The activity of restriction nucleases was assayed by digestions of λ DNA (a gift from U. Hänggi). One unit is defined as the amount required to degrade to completion 1 μ g of λ DNA in 1 h at 37° in a total assay mixture of 0.02 ml.

Isolation of DNA. DNA was isolated from purified liver nuclei (19) of 150-300g male Sprague Dawley albino rats. Nuclei containing the equivalent of 1000 A₂₆₀ units DNA were incubated

in buffer A for 30 min at 37° C in a total volume of 20 ml with 40 units (20) of preincubated (21) DNAase-free pancreatic RNAase (Worthington Biochem. Corp., Freehold NJ). RNAase treated nuclei were washed twice in buffer A containing 0.1 % Triton X-100, once in buffer A and finally suspended in 200 ml 50 mM Tris-HCl, pH 7.6, 0.1 % SDS, 1 mM EDTA and digested with 50-100 µg/ml proteinase K (E. Merck, Darmstadt) at 37° overnight or until a clear viscous solution was obtained. After these steps the procedure of Marmur (22) was followed, omitting further RNAase and protease treatments. DNA was finally dialyzed against 10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA and kept frozen at -20° C at a concentration of 5-10 A₂₆₀ units/ml. ³H-thymidine labeled DNA was prepared from nuclei as described (11).

Terminal labeling and sequence analysis. Restriction fragments were isolated from polyacrylamide gels by elution with 1 M NaCl after grinding the gel pieces and filtering through Whatman 3 MM cellulose paper (W. Fiers, pers. communication). 5' terminal labeling and sequence analysis were performed according to Maxam and Gilbert (23 and pers. communication). A fifth reaction, specific for purine residues, was employed as described by Gray et al. (24). Partial cleavage products were fractionated on 20 % or 8 % polyacrylamide, 8.3 M urea gels with dimensions of 0.25x200x400 mm. Autoradiography (Kodak X-Omat XR 5) was performed at -70° C; intensifying screens (Cawo SE 6 or Ampli Universal) were used in some experiments.

RESULTS AND DISCUSSION

Isolation of rat satellite DNA I. If a repetitive DNA component differs from the rest of the DNA by containing either a very high or alternatively a very low number of sites for a certain restriction nuclease a purification scheme can be devised based on separation by size. Botchan et al. showed that the latter approach could be applied, at least in principle, to the purification of mouse satellite DNA (25). The procedure has not been practically used since mouse satellite DNA can be isolated conveniently by density gradient centrifugation. We have now applied it to rat satellite DNA I which so far could not be purified efficiently by centrifugation techniques.

Incubations of rat DNA with Sau3AI (cleavage site GATC) converted 97-99 % of the DNA to short fragments and left satellite DNA I as a small fraction of high molecular weight (marked by arrow in track b of Fig. 1). The total amount of this fraction

Figure 1. Isolation of rat satellite DNA I by gel filtration. 100 A₂₆₀ units of rat DNA, digested with an excess of Sau3AI were loaded on a 2x100 cm column of Sepharose 2B equilibrated with 10 mM Tris-HCl, pH 7.6, 1 mM EDTA, 10 mM NaCl at 4° and eluted with the same buffer at a flow rate of 3-5 ml/h. 5-6 ml fractions were collected and after addition of NaCl to a final concentration of 0.5 M DNA was precipitated with ethanol. Individual DNA fractions were dissolved in 10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA at a concentration of 10 A₂₆₀ units/ml. 0.05 A₂₆₀ unit aliquots were submitted to electrophoreses (11) on a 1.5 % agarose gel. λ dv-1 DNA digested with HaeIII is shown in lane a as a chain length marker (26). 0.2 A₂₆₀ units of the Sau3AI digest before gel filtration is shown in (b). c-f: column fractions 14,16,20 and 30, respectively. 0.05 A₂₆₀ units of fractions 16 and 25 were digested with EcoRI and run in tracks g and h, respectively. The dotted area of the elution profile (fractions 14-16) represents fractions enriched in satellite DNA I. The arrows indicate the position of satellite I before (tracks b-d) and after (g) EcoRI digestion. The bracket in track b indicates the area where repetitive DNA different from satellite I is seen. Tracks a-f and g-h, respectively, were from two different gels.

in the rat genome was estimated to be 1-3 % by in vivo labeling experiments with ^3H -thymidine. A similar value was obtained on the basis of fluorescence intensities of fragments in agarose gels stained with ethidium bromide. Our value is comparable to the one of 2.9 % found for the HindIII monomer (13).

Satellite DNA I was purified on a preparative scale by gel filtration of a Sau3AI digest on Sepharose 2B columns (Fig. 1). Fractions 14,15, and 16 of the elution diagram of Fig. 1 contained satellite DNA I of approximately 90, 50 and 30 % purity, respectively. The satellite DNA content of the fractions was assayed by EcoRI digestion (e.g. track g of Fig. 1) and quantitative evaluation (27) of the ethidium bromide fluorescence of the resulting bands. Pooled fractions 14-16 or analogous fractions of preparations which had been upscaled by a factor of 10 contained 30-40 % satellite I and were designated partially purified satellite DNA I. Since highly repetitive components other than satellite I were present only in trace amounts, such preparations could be used without further purification in most of our work. Attempts at purifying satellite DNA I by buoyant density centrifugation in neutral CsCl gradients of undigested or Sau3AI restricted rat DNA were unsuccessful. The distribution of satellite I overlapped largely with that of the bulk DNA in those gradients. Highly purified satellite DNA I was obtained, however, from partially purified preparations by preparative gel electrophoresis and was estimated to be at least 95 % pure.

The average length of the Sau3AI fragments of the non-satellite I DNA was estimated by agarose gel electrophoresis and subsequent quantitative evaluation (14,25) of the ethidium bromide fluorescence. A number average chain length of approximately 400 bp was obtained which is slightly higher than the value expected on the basis of an average distance of 270 bp between Sau3AI cleavage sites in rat DNA assumed to have a 42 % G+C content (arithmetic mean of the values listed in ref. 28).

The average length of the satellite DNA I fragments in partially and highly purified preparations was estimated by electrophoresis on 0.3 % agarose gels to be about 10 000 bp. Smaller fragments of this satellite DNA in the 370 bp register

could not be detected when Sau3AI digests of total rat DNA were electrophoresed on 0.3-1.5 % agarose gels and stained with ethidium bromide (e.g. Fig. 1b). When the DNA of a 1.5 % gel was transferred, however, onto a nitrocellulose filter (29) and hybridized with nick-translated (30) highly purified satellite DNA I small satellite DNA fragments were seen (experiment not shown). In fact the 370 bp monomer and its lower multiples were well resolved while continuous hybridization was observed in the range of the higher multiples. This shows that part of satellite DNA I contains Sau3AI cleavage sites in the 370 bp register but analyses of the Sepharose 2B eluates (Fig. 1) by subsequent EcoRI digestion indicate that it may be only a small part.

In addition to satellite DNA I rat DNA contains at least one, probably several highly repetitive components. Sau3AI digests of rat DNA revealed a large number of sharp bands in electropherograms (e.g. track b in Fig. 1, marked by bracket) which in blotting experiments did not hybridize with nick-translated highly purified satellite DNA I. These bands may correspond to some of the fragments previously characterized in this laboratory (9,11), to the 1408 bp EcoRI and 188 bp HindIII fragments of ref. 10 and to the large fragments of ref. 12 and 30.

Restriction map and sequence analysis. The digestion and codigestion experiments with EcoRI, HindIII, HaeIII, HphI, and HinfI led to the construction of the restriction map shown in Fig. 2 which served as the basis of the sequencing work. The positions of the HinfI sites were confirmed by sequence analysis and were found not consistent with published mapping data (12) of the HindIII monomer of Novikoff hepatoma DNA.

Partially purified satellite DNA I was cleaved with HaeIII or HindIII, and dephosphorylated with alkaline phosphatase. The monomer fragments were isolated and terminally labeled with ^{32}P . After digestion with HindIII or HaeIII, respectively, the 5'-labeled fragments were used for sequence analysis according to Maxam and Gilbert (23). Thus, it was possible to sequence both strands of the 370 bp unit as illustrated in Fig. 2. In addition three HphI sites detected were used for sequencing across the HindIII and HaeIII sites (Fig. 2). The sequences of both strands

Figure 2. Restriction map and sequencing strategy of rat satellite DNA I. Restriction sites within one 370 base pair repeat unit are indicated. The direction and extent of sequence analysis is represented by solid arrows drawn below the restriction map.

could be established from the autoradiograms without major ambiguities. A representative autoradiogram is shown in Fig. 3. The sequence of 370 bp can be subdivided into two 92 bp and two 93 bp units (Fig. 4). The units were aligned to achieve maximal homology leaving a gap of 1 bp in each of the two 92 bp units. The 93 bp sequence is taken to be the basic unit of rat satellite DNA I.

The 141 GC pairs found in the 370 bp repeat unit correspond to a GC content of 38.1 %. We have no explanation for the disagreement with a previous estimate (12,13).

One problem in sequencing satellite DNA is the choice of a fraction which is as representative as possible for the whole satellite DNA. In the case of satellite DNA I from rat the purification procedure eliminates those parts which contain Sau3AI cleavage sites in close proximity. In the sequencing procedure HindIII monomers which contain one HaeIII site are selected for. This is a sizeable fraction which can be estimated to comprise about half of the satellite DNA I but we have no way to quantitate this amount exactly.

The 370 bp sequence was analysed by computer for the presence of sequences of dyad symmetry which are longer than 5 bp using a program developed by U. Hänggi. 150 bp were found to be

Figure 3. Autoradiogram of a sequencing gel of the HindIII fragment of satellite DNA I. The isolated HindIII monomer fragment of partially purified rat satellite DNA I was terminally labeled and cleaved with HaeIII. Both isolated fragments were submitted to the partial cleavage reactions (Methods). The cleavage products were fractionated on an 8 % polyacrylamide, 8.3 M urea-gel. The partial sequence of the larger (a) and smaller (b) HindIII-HaeIII fragments can be read off the autoradiogram. Asterisks indicate cytosine residues which are methylated. Nucleotide symbols set in parentheses stand for residues present in minor amounts.

involved in 6-8 bp long symmetry elements some of them being rather far apart from each other.

Comparison of the rat satellite DNA I sequence with the still incomplete sequence of mouse satellite DNA (32) revealed a rather high degree of homology (Fig. 5). Significant homologies were also seen with the satellite DNA sequences of other species. Some of the homologies could be discussed in terms of evolutionary relatedness of the species. One may also speculate that some short sequences which are found in several satellite DNAs play a role in specific protein binding and, with that, have a functional significance.

Methylation of CpG and the influence of the sequence on restriction nuclease digestion. The CpG frequency of rat satellite DNA I was found to be 2.7 % by sequence analysis and is therefore similar to the 2.6 % found in mouse satellite DNA (36) but clearly higher than the 1.2 % reported for total rat DNA (37). When m⁵C occurs in a DNA sequence a gap is found in the respective position of the sequencing gel (38,39) and a band corresponding to a G residue in the complementary strand. Very faint bands in C positions indicate that the respective C-residues are largely methylated (designated C* in Fig. 3 and 4). Methylation was detected only in the 10 CpG pairs of the sequence and was found to be symmetrical in all cases.

Cleavage by restriction nucleases is known to be influenced by the methylation pattern of the recognition site. The
 (C*) GGCCC sequence of rat satellite DNA was found to be cleaved
 (G) CCGGG
 by HaeIII but only very slowly by Sau96I. The cleavage by HaeIII is in agreement with the results of ref. (40). Resistance against

Figure 4. Sequence of the 370 bp repeat unit of rat satellite DNA I. Nucleotides occurring in identical positions in the four 92/93 bp units are emphasized by shading. Restriction nuclease cleavage sites are indicated by brackets. Asterisks indicate methylation and nucleotide symbols set in small type stand for residues present in minor amounts.

Figure 5. Homologies between the sequences of rat satellite DNA I (III) and satellite DNA sequences of D. melanogaster (I), mouse (II), calf (IV), man (V) and monkey (VI). The following satellite DNA sequences are aligned: part of the sequence of the 1.688 D. melanogaster satellite (33), the still incomplete "composite sequence" of mouse satellite DNA (32), the sequence from position 310 to 34 of rat satellite DNA I (this paper), the prototype sequence of the Sau segments in the 1.706 bovine satellite (M. Pech, R.E. Strebeck, and H.G. Zachau, submitted for publication), positions 61-88 of a human satellite DNA sequence (34), and positions 39-66 of the C. aethiops alpha-satellite sequence (35). Homologies are emphasized by shading.

Sau96I had previously been observed in the related sequence
GGCCC*(^G_{C*}) of the 1.706 bovine satellite DNA (M. Pech, R.E. Streeck,
and H.G. Zachau, manuscript submitted).

Methylation of C-residues in EcoRI and HinfI sites of rat satellite DNA I (Fig. 4) appears to influence the rate of cleavage by these nucleases (Fig. 6). In EcoRI digests the 370 bp band and its integral multiples are more intensive than the other bands (track b-d in Fig. 6). The same intensity distribution was obtained with unfractionated rat DNA (not shown) regardless of whether the enzyme concentrations or the incubation times were varied. The intensity differences were maintained also in blotting and hybridization experiments. The rapidly cleaved EcoRI site was localized by codigestion experiments with HindIII and HaeIII at positions 1-6 (Fig. 4) which had been found to be unmethylated. 2 of the 3 slowly cleaved EcoRI sites are methylated (positions 186-191 and 279-284). We have no explanation for the observation that one unmethylated site (positions 94-99) is split more slowly than the other unmethylated site (positions 1-6) except for saying that outside sequences somehow must influence the rate of cleavage.

The intensity distribution in the HinfI pattern (Fig. 6) is more complex than in the EcoRI pattern in agreement with the non-equidistant distribution of the HinfI sites. Using a quantitative relationship similar to the one of ref. 41 the intensity of the bands and, with that, the rate of cleavage can be correlated with the methylation pattern: the methylated sites at positions 73-77 and 298-302 were cleaved slowly, while the unmethylated sites at positions 113-117 and 205-209 were cleaved fast; the methylation of the adjacent C (position 118) seems to influence the rate of cleavage only little or not at all.

Divergence of satellite DNA I sequence and evolutionary aspects. Two results of the sequence analysis and two observations in the restriction nuclease analysis of satellite DNA I can be related to sequence divergence. The four points are described in the following and then discussed together in terms of an evolutionary model.

(a) A prototype sequence can be derived by selecting the

Figure 6. Course of digestion of satellite DNA I with EcoRI and HinfI. 1.25 μ g each of satellite I were digested with 16,40,80, 160,800 and 8000 U/ml EcoRI, respectively, in a total volume of 25 μ l incubation mixture at 37 $^{\circ}$ for 1 h and submitted for electrophoresis (a-f) on a 1.5 % agarose gel. Digestions with HinfI were done under the same conditions with 3,6,12,20 and 240 units/ml, respectively, (g-l). The DNA in (l) was digested for 5 h instead of 1 h. The material in (m) is the same as in (d). The numbers indicate the monomer and integral multiples of the 92/93 bp unit.

nucleotides occurring most frequently in homologous positions of the four aligned 92/93 bp units (Fig. 4). The overall divergence of these units from the prototype sequence amounts to 19 %. This divergence is calculated as the number of nucleotide exchanges in the actual sequence compared with the prototype sequence. Comparing individual 92/93 bp units to each other instead of comparing them to the prototype sequence leads to homologies of 62-73 %. Only two regions larger than 3 bp are identical in all 4 units (shaded in Fig. 4), that is the sequence TGGGAAC and the EcoRI restriction site GAATTC.

(b) At defined positions in the autoradiograms faint bands appeared in addition to the major ones (Fig. 3). Complementary "minor" nucleotides were found in the other strand. With that and by their intensities the faint bands could be clearly distinguished from those bands resulting from non-specific cleavage. The faint bands represent small amounts of other nucleotide residues at the respective positions of the sequence. Those nucleotides are shown in Fig. 3 in parentheses and in Fig. 4 in small type. Apparently some satellite DNA fragments have a sequence different from the majority of fragments. The differences are confined to certain positions and only to one alternative nucleotide. Whether the alternative nucleotides, which were found in 8 positions, can be ascribed to only one additional population of 370 bp repeat units or whether each one is represented in a different population cannot be decided by the present techniques. In general, nucleotide exchanges cannot be detected by the sequencing procedure if they are present in less than 5-10 % of the repeat units. The nucleotide sequence, shown in Fig. 4, represents therefore the sequence of the most frequent nucleotides at each single position in the 370 bp unit.

(c) A general phenomenon observed in restriction nuclease cleavage of satellite DNAs is also seen with satellite DNA I. If EcoRI and HaeIII sites found in the satellite DNA I sequence (Fig. 4) occurred in all units one should find upon extensive digestion only monomer fragments. But this was not the case. Complete EcoRI digests comprise only 73 % of the

monomer and, in addition, 21 % dimer, 4 % trimer, and 1 % tetramer. In HaeIII digests 91 % monomer, 7 % dimer, and 1 % trimer were found. In these experiments the ethidium bromide fluorescence intensities were evaluated. Trace amounts of higher bands could be seen in blotting and hybridization experiments. The degree of divergence was calculated from the ratio of the amounts of the fragments (41,42) in the limit digests. Assuming random mutations within the cleavage sites one arrives at values of 3.2 % and 3.1 % for the EcoRI and HaeIII sites, respectively.

(d) It should be mentioned that digestion of satellite DNA I with EcoRI, HaeIII, or HindIII left approximately 15 % of the material at a high molecular weight, although the digestions were complete by the usual criteria. Most of this material could be cleaved by one of the other restriction nucleases to yield a regular pattern; the material therefore consisted of satellite DNA I lacking the cleavage sites for the respective nuclease.

On the basis of the sequence comparisons (a) one may assume that the 370 bp sequence was formed from a 93 bp unit relatively early in evolution; the differences between the four 92/93 bp units within the 370 bp sequence were introduced by mutation, which gave rise also to HindIII and HaeIII sites in the 370 bp register. The present day satellite DNA I would then have been formed from the 370 bp sequence by amplification and additional mutations (b-d).

The results (b) and (d) may be related to a particular feature of the amplification process. The observation of a defined small number of nucleotide substitutions and the finding of small fractions of the satellite DNA being completely devoid of certain restriction sites can be explained best by the following scheme: at an early stage of the amplification process different mutational events have taken place in different copies of the 370 bp sequence which were then amplified to differing degrees giving rise to fractions which are distinguished by the presence or absence of certain nucleotides at certain cleavage sites. Comparable fractions have been found first in mouse satellite DNA and were called segments (43). A similar segmentalization was also observed in the satellite DNAs of several

other organisms (reviews 2,44) and was, in the case of human satellite DNA III, clearly shown to be related to its localization on different chromosomes (45).

ACKNOWLEDGEMENT

We are indebted to U. Ziegler and I. Brenner for expert assistance. The work was supported by Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

REFERENCES

1. Skinner, D.M. (1977) *Bioscience* 27, 790-796.
2. Appels, R. and Peacock, W.J. (1978) *International Review of Cytology*, suppl. 8. Bourne, G.H. and Danielli, J.F. eds. pp. 69-126. Acad. Press.
3. Smith, G.P. (1978) *Trends in Biochem. Sci.* 3, N34-36.
4. John, B. and Miklos, G.L.G. (1979). *International Review of Cytology* 58, in press.
5. Bonner, J., Garrard, W.T., Gottesfeld, J., Holmes, D.S., Sevall, J.S., and Wilkes, M. (1973). *Cold Spring Harbor Symp. Quant. Biol.* 38, 303-310.
6. Walker, P.M.B. (1968) *Nature* 219, 228-232.
7. Roizes, G. (1974) *Nucl. Acids Res.* 1, 1099-1120.
8. Guttman, T., Votavova, H., and Pivec, L. (1976) *Nucl. Acids Res.* 3, 835-845.
9. Philippsen, P., Streeck, R.E., and Zachau, H.G. (1974) *Eur. J. Biochem.* 45, 479-488.
10. Maio, J.J., Brown, F.L., and Musich, P.R. (1977) *J. Mol. Biol.* 117, 637-655.
11. Igo-Kemenes, T., Greil, W., and Zachau, H.G. (1977) *Nucl. Acids Res.* 4, 3387-3400.
12. Fuke, M. and Busch, H. (1979) *FEBS Lett.* 99, 136-140.
13. Fuke, M., Davis, F.M., and Busch, H. (1979) *FEBS Lett.* 102, 46-50.
14. Igo-Kemenes, T. and Zachau, H.G. (1977) *Cold Spring Harbor Symp. Quant. Biol.* 42, 109-118.
15. Igo-Kemenes, T., Miller, F., and Zachau, H.G. (1978). *FEBS 12th Meeting Dresden 1978, Vol. 51, Gene Functions*, pp. 219-229. S. Rosenthal et al. eds. Pergamon Press Oxford and New York 1979.
16. De Filippos, F.M. (1974) *Biochem. Biophys. Res. Commun.* 58, 586-596.
17. Sussenbach, J.S., Monfoort, C.H., Schiphof, R., and Stobberingh, E.E. (1976) *Nucl. Acids Res.* 3, 3193-3202.
18. Sussenbach, J.S., Steenbergh, P.H., Rost, J.A., van Leeuwen, W.J., and van Embden, J.D.A. (1978) *Nucl. Acids Res.* 5, 1153-1163.
19. Hewish, D.R. and Burgoyne, L.A. (1973) *Biochem. Biophys. Res. Commun.* 52, 504-510.
20. Kunitz, M. (1940) *J. Gen. Physiol.* 24, 15-32.

21. Meyers, J.A., Sanchez, D., Elwell, L.P., and Falkow, S. (1976) *J. Bacteriol.* 127, 1529-1537.
22. Marmur, J. (1961) *J. Mol. Biol.* 3, 208-218.
23. Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. US* 74, 560-564.
24. Gray, C.P., Sommer, R., Polke, C., Beck, E., and Schaller, H. (1978) *Proc. Natl. Acad. Sci. US* 75, 50-53.
25. Botchan, M., McKenna, G., and Sharp, P.A. (1973) *Cold Spring Harbor Symp. Quant. Biol.* 38, 383-395.
26. Streeck, R.E. and Zachau, H.G. (1978) *Eur. J. Biochem.* 89, 267-279.
27. Pulleyblank, D.E., Shure, M., and Vinograd, J. (1977) *Nucl. Acids Res.* 4, 1409-1418.
28. *Handbook of Biochemistry* (1970) H.A. Sober ed. CRC Press Cleveland, pp. H-97.
29. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.
30. Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.
31. Botchan, M., Topp, W., and Sambrook, J. (1976) *Cell* 9, 269-287.
32. Shmookler-Reis, R.J. and Biro, P.A. (1978) *J. Mol. Biol.* 121, 357-374.
33. Brutlag, D., Carlson, M., Fry, K., and Hsieh, T.S. (1978) *Cold Spring Harbor Symp. Quant. Biol.* 42, 1137-1146.
34. Manuelidis, L. and Wu, J.C. (1978) *Nature* 276, 92-94.
35. Rosenberg, H., Singer, M., and Rosenberg, M. (1978) *Science* 200, 394-402.
36. Russell, G.J., Walker, P.M.B., Elton, R.A., and Subak-Sharpe, J.H. (1976) *J. Mol. Biol.* 108, 1-23.
37. Setlow, P. (1975) in *Handbook of Biochemistry and Molecular Biology*, 3rd Edition, Nucleic Acids Vol. II (Fasman, G.D. ed.) pp. 312-318, CRC Press Cleveland.
38. Ohmori, H., Tomizawa, J., and Maxam, A.M. (1978) *Nucl. Acids Res.* 5, 1479-1485.
39. Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V., and Brown, D.D. (1978) *Cell* 13, 717-725.
40. Mann, M.B. and Smith, H.O. (1977) *Nucl. Acids Res.* 4, 4211-4221.
41. Southern, E.M. (1975) *J. Mol. Biol.* 94, 51-69.
42. Hörz, W. and Zachau, H.G. (1975) in *Biochemistry of the Cell Nucleus and Mechanism and Regulation of Gene Expression* (E.J. Hidvegi, J. Sumegi, and P. Venetianer, eds.) pp. 403-408, North-Holland Amsterdam.
43. Hörz, W. and Zachau, H.G. (1977) *Eur. J. Biochem.* 73, 383-392.
44. Dover, G. (1978) *Nature* 272, 123-124.
45. Beauchamp, R.S., Mitchell, A.R., Buckland, R.A., and Bostock, C.J. (1979) *Chromosoma* 71, 153-166.