Volume 5 Number 3 March 1978 Nucleic Acids Research

Recombinant plasmids containing Xenopus laevis globin structural genes derived from complementary DNA

Peter Humphries, Robert Old, Lesley W. Coggins, Theresa McShane, Christine Watson and John Paul

Beatson Institute for Cancer Research, Wolfson Laboratory for Molecular Pathology, Garscube Estate, Switchback Road, Glasgow G61, UK

Received 12 January 1978

ABSTRACT

Details are presented of the *in vitro* synthesis of double-stranded DNA complementary to purified *Xenopus* globin messenger RNA, using a combination of reverse transcriptase, fragment 'A' of *E. coli* DNA polymerase 1 and S1 endonuclease. After selection of duplex DNA molecules approaching the length of *Xenopus* globin messenger RNA by sedimentation of the DNA through neutral sucrose gradients, the 3'-OH termini of the synthetic globin gene sequences were extended with short tracts of oligo dGMP using terminal transferase. This material was integrated into oligo dCMP-extended linear pCR1 plasmid DNA and amplified by transfection of *E. coli*. Plasmids carrying globin sequences were identified by hybridization of ^{32}P-labelled globin mRNA to total cellular DNA *in situ*, by hybridization of purified plasmids to globin cDNA in solution, by analysis of recombinant DNA on polyacrylamide and agarose gels, and by heteroduplex mapping. The results show that extensive DNA copies of *Xenopus* globin mRNA have been integrated into recombinant plasmids.

INTRODUCTION

Several methods are currently available for the construction of recombinants between plasmid or phage vectors and eukaryotic structural genes. These methods involve either the direct ligation of restriction fragments of *in vitro* synthesized gene sequences to a vector carrying the same restricted termini (1,2), or the actual construction of cohesive termini involving dG-dC or dA-dT linkages generated by terminal transferase (3-6, 29).

The structural organization of amphibian genomes has been the subject of intensive study (reviews, refs. 7,8) and the availability of hybridization probes of absolute sequence purity would be very valuable, particularly in studies of transcription in amphibian oocytes (9).

We report here the results of experiments designed to construct hybrid plasmids carrying *Xenopus laevis* globin structural gene sequences. The procedure does not require the filling in or removal of protruding 5'-termini of plasmid vector DNA prior to the addition of cohesive homopolymer tails and is sufficiently sensitive to be applicable to the transfection

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

and amplification in E. coli of as little as one nanogram of double-stranded cDNA.

MATERIALS AND METHODS

1. Enzymes.

AMV[a] reverse transcriptase was purified according to the procedure of Kacian and Spiegelman (10). 1 unit was defined as that quantity of enzyme that produced a 10% yield (i.e. 6μg) of cDNA from 60μg of globin mRNA in 1ml of a reaction mixture containing 2mM each of dCTP, dATP, dTTP and dGTP, 50mM tris-HCl pH 8.2, 10mM magnesium acetate, 2mM dithiothreitol, 40mM NaCl, 75μg actinomycin D and 40μg oligo dT after incubation for 2 hours at 37°C. Terminal deoxynucleotidyl transferase was purified from foetal calf thymus glands according to a modified procedure of Bollum (11). 1 unit was defined as that quantity of enzyme that incorporated 1 pmol dGMP into acid insoluble product in 10 minutes at 37°C in a mix containing 10mM Hepes-NaOH buffer pH 7.2, 4mM $MgCl_2$, 1mM 2-mercaptoethanol, 240μM dGTP and 3μg oligo dA12-18. E. coli DNA polymerase 1 fragment 'A' (12) was purchased from the Boehringer Corporation. The unit of activity was defined according to Klenow, et al.(12). Restriction endonuclease Eco R1 was prepared according to the procedure of Green et al (13). The optimum quantity of enzyme required to linearise pCR1 plasmid DNA was determined by analysis of the DNA by gel electrophoresis and under the electron microscope. Restriction endonucleases Hha, Hae III and Hind III were purchased from the New England Biolab Company. 1 unit was that quantity of enzyme required to digest 1 microgram of lambda phage DNA in 1 hour at 37°C. Endonuclease S1 from Aspergillus oryzae, purified according to the procedure of Vogt (14) was the gift of Dr. Mike Houghton and colleagues of Searle Research Laboratories. 1 unit was that quantity of enzyme that digested 1 microgram of single stranded DNA in 1 hour at 40°C. Polynucleotide kinase was purchased from PL Biochemicals. The unit of enzyme activity was defined according to Richardson (28).

2. Preparation of globin mRNA from X. laevis.

Induction of anaemia.

Preparations were usually performed with 5 or 10 large (greater than 8cms long from mouth to anus) mature female X. laevis. Animals were anaesthetized by immersion in 0.2% MS222[a] and injected subcutaneously with 0.5ml of 0.5% phenylhydrazine hydrochloride. This was repeated 48 hours later. Between two to four weeks after the first injection the blood of

treated animals showed a low concentration of cells (10-50% normal values) with a high proportion of immature red cells, as described by Thomas and Maclean (15). Blood was collected from anaesthetized anaemic animals by ventricular puncture (typically 3-5ml blood was obtained from each individual). The blood was delivered into ice-cold NTM solution (0.14M NaCl, 0.01M tris-HCl pH 7.4, 0.003M $MgCl_2$) containing 25 units/ml of heparin.

Isolation of mRNA

Red blood cells were washed twice with NTM at 4°C. The packed washed cells (1 volume) were resuspended in 3 volumes of cold NTM and were lysed by the addition of a further 3 volumes of NTM containing 1% Nonidet P40 (Shell), 0.5M sucrose, followed by gently agitation. (Vigorous agitation resulted in lysis of the sensitive red cell nuclei). The lysate was centrifuged at 13,000rpm for 15 minutes in a Sorvall HB4 rotor at 4°C. SDS (10%) and EDTA (0.4M) were added to the cleared lysate to give final concentrations of 0.5% and 10mM respectively.

Total cytoplasmic RNA was prepared by 4 extractions with phenol-chloroform-isoamyl alcohol (50:50:1). RNA was recovered by ethanol precipitation. The RNA was resuspended in 0.1M NaCl, 0.01M tris-HCl pH 7.4, 0.001M EDTA for chromatography on a poly (U)-Sepharose column as described (16). Polyadenylated mRNA eluted from the column was precipitated with ethanol and stored at -20°C. The yield of mRNA from 5 animals was about 50µg.

3. Translation of mRNA.

Messenger RNA was translated in a cell-free system prepared from wheat germ. Reactions were performed as described by Roberts and Paterson (17) with modifications. Each reaction contained, in 25µl; 10µl wheat germ extract, 1µg mRNA (as appropriate), 1mM ATP, 0.2mM GTP, 2mM dithiothreitol, 0.5mM spermine, 0.25mM spermidine, 16mM creatine phosphate, 40µg/ml creatine phosphokinase (Sigma Biochemicals), 24mM Hepes-KOH buffer pH 7.2, 1.5mM magnesium acetate, 92mM KCl, 2.4mM 2-mercaptoethanol, 200 µM each unlabelled amino acid minus methionine, and 6.5µCi L-(35S) methionine (Radiochemical Centre, Amersham: specific radioactivity 500 Ci/mmol).

Reactions were incubated at 23°C for 90 minutes. Samples from the translation reaction were analysed by SDS-polyacrylamide gel electrophoresis as described by Laemmli (18). After electrophoresis gels were stained with Coomassie Brilliant Blue, destained in 7.5% acetic acid, dried under reduced pressure, and autoradiographed on X-ray film (Fuji Rx).

4. Synthesis of oligo dGMP-extended duplex globin cDNA.

cDNA was synthesized in 1.0ml of reaction mixture containing 25µg

globin mRNA, 2mM each of dATP, dGTP, dTTP and ^{3}H-dCTP (125μCi/μmol), 50mM tris-HCl pH 8.2, 10mM magnesium acetate, 2mM dithiothreitol, 40mM NaCl, 40μg oligo dT, 75μg actinomycin D, 375μg bovine serum albumin and 1 unit of reverse transcriptase. After incubation for 2 hours at 37°C the mixture was extracted with phenol and passed through a column of Sephadex G-50. Fractions containing the cDNA were pooled, precipitated with ethanol and re-dissolved in 300μl of 0.3M NaOH. After incubation for 30 minutes at 60°C, the cDNA was again precipitated with ethanol and re-dissolved in 100μl of water.

2μg of single-stranded globin cDNA was then incubated in 420μl of reaction mixture containing 30mM tris-HCl pH 7.5, 4mM $MgCl_2$, 0.5mM 2-mercaptoethanol, 1mM each of dATP, dCTP, dGTP and dTTP, and 9 units of E. coli DNA polymerase 1 fragment 'A'. After incubation for 5 hours at 22°C the mixture was extracted with phenol, dialysed extensively against water and lyophilized to dryness. The partially double-stranded cDNA was then re-dissolved in 32μl of a reaction mixture containing 100mM NaCl, 50mM sodium acetate pH 4.5, 1mM zinc sulphate, and 5 units of endonuclease S1. After incubation for 2 hours at 43°C the reaction was extracted with phenol, dialysed extensively against 1mM tris-HCl pH 7.5 and lyophilised to dryness. After re-suspension in 100μl of 10mM tris-HCl pH 7.5, 0.1M NaCl the duplex cDNA was layered onto the surface of a 5-20% (w/v) neutral sucrose gradient and centrifuged for 20 hours at 20°C and at 110,000xg in order to select those DNA molecules of a size approaching that of full-length X. laevis globin mRNA.

Then 40ng of 'full-length' ds-cDNA[a] was suspended in 300μl of a reaction mixture containing 5mM $MgCl_2$, 1mM 2-mercaptoethanol, 0.6mM dGTP and 12.5mM Hepes-NaOH buffer pH 7.1. 1000 units of terminal transferase was added, and after incubation for 5 minutes at 37°C the reaction was extracted with phenol, dialysed extensively against 1mM tris-HCl pH 7.3, lyophilized to dryness and stored at -20°C.

5. Linearization and oligo dCMP-extension of pCR1 DNA.

Superhelical pCR1 plasmid DNA was purified by equilibrium banding in caesium chloride followed by sedimentation of the DNA through neutral 5-20% sucrose gradients (19). 100μg of plasmid DNA was then incubated in 200μl of reaction mixture containing 10mM NaCl, 8mM tris-HCl pH 7.5, 2mM $MgCl_2$, 2mM 2-mercaptoethanol and a sufficient quantity of Eco R1 restriction endonuclease to render all of the DNA molecules linear after incubation for 30 minutes at 37°C. The linearization reaction was monitored by analysis

of the plasmid DNA on 1% agarose gels and under the electron microscope. The reaction mixture was then extracted with phenol, the DNA precipitated with ethanol and re-dissolved in 400µl of 4mM tris-HCl pH 7.9. After extensive dialysis against the same buffer, 40µg of linear plasmid DNA was incubated at 37°C in 1.8ml of a reaction mixture containing 200mM Hepes-NaOH pH 7.1, 1mM $CoCl_2$, 1mM 2-mercaptoethanol, 1.25mM dCTP and 6000 units of terminal transferase. Indubation for 5 minutes resulted in the addition of an average of 30 residues of dCMP to each 3'-OH terminus of the plasmid DNA. The reaction mixture was extracted with phenol, precipitated with ethanol and re-dissolved in 300µl of 2mM tris-HCl pH 7.3. After extensive dialysis against the same buffer, the 'tailed' linear plasmid DNA was stored at 4°C.

6. Formation of hybrids between 'tailed' plasmid and globin DNA sequences.

150ng of linear, oligo dCMP extended pCR1 plasmid DNA and 20ng of oligo dGMP extended duplex globin cDNA were incubated in 200µl of reaction mixture containing 100mM NaCl, 10mM tris-HCl pH 7.5 and 1mM EDTA. After heating for 1 hour at 65°C the mixture was allowed to cool for 2 hours to room temperature. Then, 300µl of 100mM tris-HCl pH 7.3 was added, followed by 1.2ml of calcium-treated competent cells of E. coli C600 r_k^- m_k^- rec B^-C^- thy^-. The bacteria were prepared for transformation as previously described (6). The mixture was maintained on ice for 10 minutes, followed by 5 minutes incubation at 37°C. Then 2.5ml of L-broth was added and the transfected bacterial suspension was incubated for 1 hour at 37°C. 200µl samples were then spread over the surface of agar plates containing 40µg/ml kanamycin sulphate and incubated at 37°C.

7. Polynucleotide kinase reaction.

X. laevis globin mRNA was radiolabelled for colony hybridization in a reaction catalysed by T4 polynucleotide kinase. In order to provide 5'-OH termini for the labelling reaction, mRNA was partially degraded by heating at 68°C for 20 minutes in 10mM tris-HCl pH 8.8. The terminal labelling reaction (25µl) contained: 1.5µg RNA, 2.5µM unlabelled ATP, 40µCi gamma-^{32}P-ATP (Radiochemical Centre, Amersham; specific radioactivity, 3000 Ci/mmol), 50mM tris-HCl pH 7.5, 10mM $MgCl_2$, 10mM dithiothreitol and 3.5 units of polynucleotide kinase. After incubation at 37°C for 2 hours, the reaction was stopped by the addition of SDS and EDTA to 1% and 20mM respectively. Yeast RNA (PL Biochemicals), 50µg was added as carrier. RNA was chromatographed on a column of Sephadex G50 in 20mM tris-HCl pH 7.4. The excluded peak of radioactive RNA (2.5×10^6 Cerenkov cpm) was recovered by ethanol precipitation.

8. Colony hybridization on nitrocellulose filters.

Preparation of filters.

Bacterial colonies containing DNA sequences complementary to _X. laevis_ globin mRNA were detected by colony hybridization (20). Millipore HAWP nitrocellulose sheet was cut into circles (85mm diameter), autoclaved at 121°C for 15 minutes and placed on the agar in 90mm petri-dishes (containing L-agar plus 40μg/ml kanamycin). Bacteria were transferred to the filter surface by streaking with sterile toothpicks. Up to 54 colonies could be conveniently accommodated in a grid-like pattern on each filter. At the same time, a reference set of colonies was prepared. The plates were incubated overnight at 37°C.

To prevent movement of bacteria and DNA during lysis and subsequent operations, solutions were applied to the underside of the filters. Whatman 3MM paper was placed on a glass plate and wetted with the appropriate solution. Filters were placed on the glistening wet surface, whereupon the solution diffused rapidly to the colonies. Between treatments filters were blotted upon dry 3MM paper.

The filters bearing bacterial colonies produced by incubation overnight were first placed upon 3MM paper wetted with 0.5N NaOH, and left in position for 10 minutes. This step lysed the bacteria and denatured the DNA. The filters were then treated in the same manner with 1.0M tris-HCl pH 7.4 (3 times), and finally with 1.5M NaCl, 0.5M tris-HCl pH 7.4. Individual filters were then placed in petri-dishes and 5ml of a solution of proteinase K (1mg/ml) in 1xSSC was added. After incubation for 1 hour at room temperature the filters were rinsed with 2xSSC, dried and baked at 80°C for 2 hours.

Hybridization.

The baked filters were pre-treated at 68°C for 5-20 hours in a solution of 3xSSC supplemented with 0.02% Ficoll, 0.02% polyvinyl pyrollidone, 0.02% bovine serum albumin (21). Following pre-treatment, hybridization reactions were performed with ^{32}P-labelled globin mRNA (10^6-10^7 Cerenkov cpm) in 10ml of 3xSSC supplemented as in the pre-treatment and also containing 0.5% SDS. These reactions were conveniently performed by sealing the filters and hybridization solution in a polythene sac, which was then immersed in a water-bath at 68°C for 16-24 hours. Filters were then removed, washed at 60°C for 3 hours in several changes of 2xSSC, 0.5% SDS, dried in air, and exposed at -70°C to X-ray film (Fuji Rx) combined with an image intensification screen.

9. Solution hybridization of globin cDNA to excess recombinant plasmid DNA

Plasmid DNA excess hybridization was carried out by hybridizing 10μg

of sonicated purified recombinant plasmid DNA with 0.5ng ^{3}H-labelled globin cDNA (specific activity $1.4x10^{7}$ cpm/µg) at 68°C for 2 hours. The incubation Cot was 2.4. The extent of cDNA hybridization was assayed by resistance to endonuclease S1 as described by Birnie et al (22).

10. Heteroduplex mapping.

Hind III-restricted recombinant plasmids (C13 or B52) and Hind III-restricted pCR11 (0.1µg each) in 20µl 50% formamide, 0.1M tris-HCl pH 8.5, 0.01M EDTA were heated to 70°C for 1 minute in a sealed Repelcoted micro-capillary to denature the DNA, and incubated at 37°C for 5 hours. 40µl of water, 30µl of formamide and 10µl of cytochrome C (0.1µg/ml) were added and the mixture was spread onto 10% formamide, 0.01M tris-HCl pH 8.5, 1mM EDTA (23). Samples were picked up on collodion coated grids, stained with uranyl acetate and shadowed with platinum-palladium (80:20).

RESULTS

1. Translation of purified X. laevis globin mRNA.

Globin mRNA was efficiently translated in the wheat germ cell-free system. Products of translation were separated on polyacrylamide slab gels in the presence of SDS. The autoradiograph (fig. 1) shows that the mRNA gave rise to a single major protein band that co-migrated with authentic X. laevis globin. The mobility relative to marker proteins (not shown) indicated an apparent molecular weight of 14,000.

2. Synthesis of oligo dGMP-extended duplex globin cDNA.

Single-stranded globin cDNA was analysed on formamide-containing 3% polyacrylamide gels as described in fig. 2. While occasionally some low molecular weight material was detectable between gel fractions 30-45 (fig. 2) the bulk of the cDNA normally ran as a sharp band having the same mobility as X. laevis globin mRNA with an apparent length of about 550 nucleotides by comparison with mouse 9S RNA.

As previously observed (6) the cDNA was able to serve as a template for fragment 'A' of E. coli DNA polymerase 1, without the addition of a primer. Fig. 3 illustrates the kinetics of incorporation of nucleotides into the second cDNA strand as assayed by resistance of the parental strand to S1 endonuclease. Normally 30-45% of the cDNA was made double-stranded after 5 hours of incubation at 22°C. Increasing the temperature and time of incubation did not significantly alter the yields of ds-cDNA. After treatment with S1 endonuclease to remove any single-stranded regions together with the

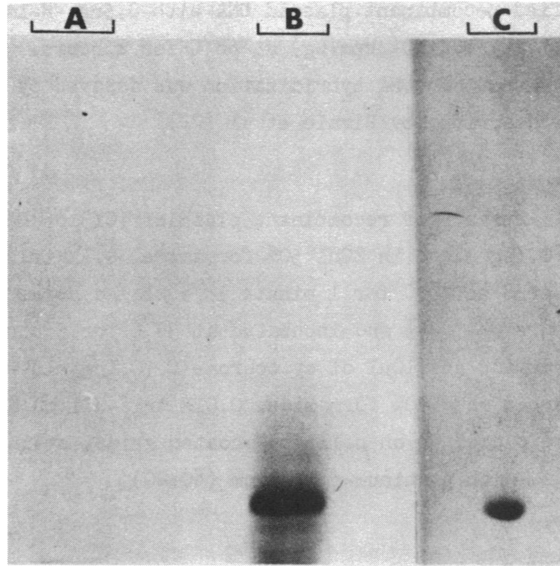

Figure 1 Autoradiograph of ^{35}S-methionine labelled X. laevis globin polypeptides synthesised in a wheat germ cell-free system and separated on a SDS-15% polyacrylamide slab gel. Samples contained: (A) no added RNA (B) 1μg of X. laevis globin mRNA. Lane (C) shows radiolabelled globin polypeptides prepared by incubation of anaemic red blood cells with ^{35}S-methionine essentially as described (15).

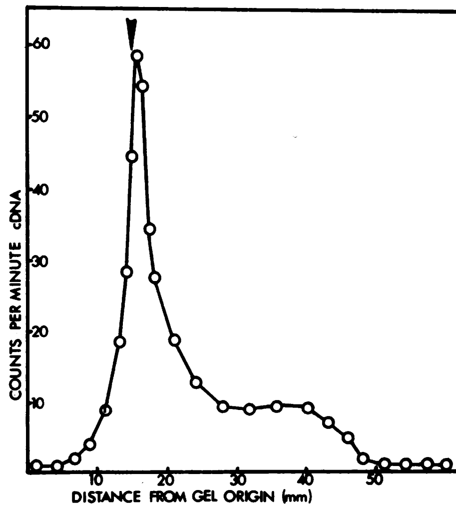

Figure 2 Analysis of X. laevis globin cDNA on formamide-containing 3% acrylamide gels. 200ng of cDNA (10,000 cpm) was run on a gel 0.5cms diameter x 10cms, for 1 hour at 5mA. The gel was sliced into 2mm segments, which were then dissolved by treatment with 1M perchloric acid and counted. The arrow indicates the point of migration of mouse 9S RNA run under similar conditions.

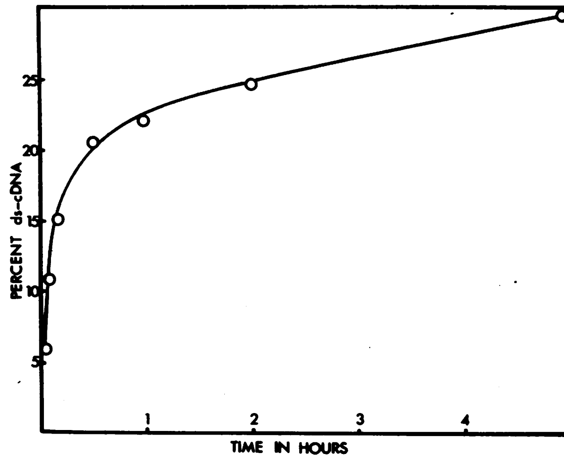

Figure 3 Kinetics of ds-cDNA synthesis. Reaction conditions are as described in the text. At time intervals indicated, aliquots were removed from the reaction mixture, treated with S1 endonuclease as described and the ds-cDNA precipitated with trichloroacetic acid and counted.

terminal hairpin at one end of the duplex DNA molecules (4,6), the ds-cDNA was sedimented through a neutral 5-20% sucrose gradient in order to select DNA molecules of a size approaching that of full length X. laevis globin mRNA. The DNA was found to sediment as a rather broad peak (fig. 4) of average length approximately 550 nucleotide pairs as determined from the S-value of the sedimenting DNA using the equations of Studier (24).

In order to select the longest ds-cDNA molecules, only the fastest-sedimenting portion of the DNA peak was selected as shown (fig. 4). This material was then 'tailed' with oligo dGMP using terminal transferase.

3. Homopolymer extension of linear pCR1 plasmid and duplex globin cDNAs.

Terminal transferase was purified to virtual homogeneity from foetal calf thymus glands according to the procedure of Bollum (11) with some modifications. The enzyme contained low levels of an endonucleolytic activity which slowly converted superhelical pCR1 DNA into its open circular form, as assayed by analysis of the plasmid DNA on 1% agarose gels (not shown). However, this activity did not cause spurious homopolymer addition to plasmid DNA. As illustrated in fig. 5, no incorporation of dCMP into superhelical plasmid could be detected under conditions in which an average of 30 residues of dCMP was added to linear pCR1 DNA. In addition, using the present reaction conditions in which cobalt ions are substituted for magnesium (25), the linear plasmid molecules served as an excellent primer for terminal

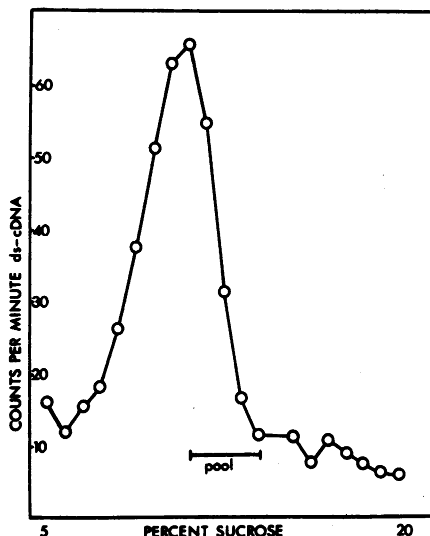

Figure 4 Neutral sucrose gradient sedimentation of ds-cDNA . Approximately 100ng of ds-cDNA was sedimented through a neutral 5-20% sucrose gradient as described in Materials and Methods. The fastest-sedimenting portion of the cDNA peak was pooled, as shown.

transferase in spite of the protruding 5'-terminal cohesive ends generated by the Eco Rl restriction enzyme.

Addition of oligo dGMP to the 3'-OH termini of ds-cDNA was monitored directly by transfection experiments involving hybrids between dGMP-tailed cDNA and linear pCRl plasmid DNA to which had been added an average of 30 residues of oligo dCMP to each 3'-OH terminus (fig. 5). Incubation of 40ng of ds-cDNA with 1000 units of terminal transferase for 5 minutes, as described in Materials and Methods produced the greatest number of transformants (see below). cDNA molecules extended using smaller quantities of terminal transferase produced proportionately fewer recombinant plasmids.

4. Amplification in E. coli of recombinant plasmids.

The efficiency of formation of viable molecular hybrids between oligo dG-extended ds-cDNA and oligo dC- extended pCRl DNA varied considerably depending on the molar ratios of both species of DNA in the hybridization mix. For example, hybrid plasmids formed between 20ng of tailed cDNA and 500ng of tailed pCRl DNA (i.e. a 1:1 plasmid to cDNA molar ratio) gave rise

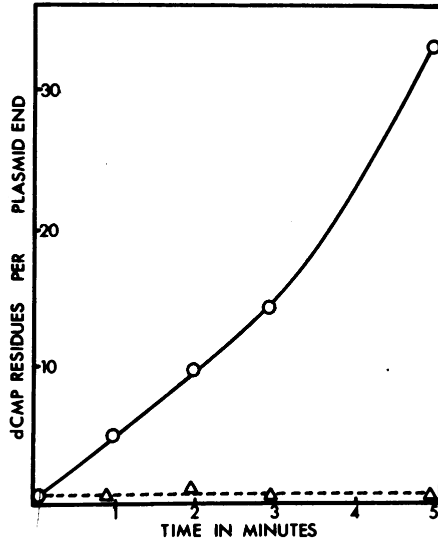

Figure 5 Activity of terminal transferase on Eco R1-linear, and superhelical pCR1 DNA. Reaction conditions are described in Materials and Methods. At 1 minute intervals, 10µl aliquots were removed from the reaction, and incorporation of ^{3}H-dCMP into plasmid DNA was assayed by TCA precipitation. o——o , linear pCR1; Δ---Δ, superhelical pCR1.

to approximately 50 recombinant clones. However, hybrids formed in the presence of a 3-fold molar excess of plasmid DNA (i.e. 20ng cDNA and 1500ng of plasmid) gave rise to approximately 200 transfectants. In control experiments, 1500ng of tailed plasmid DNA alone, produced on average 15 transfectants.

In order therefore to obtain the greatest number of recombinants it is advantageous to determine with accuracy the size of the duplex cDNA. This can be conveniently done by sucrose gradient sedimentation as described above (fig. 4).

5. Detection of recombinants containing globin DNA sequences by colony hybridization.

According to the present experimental design the only means by which a plasmid can circularize and hence become infective is by combining with tailed globin cDNA. In practice, the number of globin-carrying plasmids is somewhat less than 100% of the total recombinants. This is due firstly, to the fact that oligo dC-tailed linear plasmids alone are able to transfect E. coli at about 7% of the efficiency of globin-containing hybrid plasmids (see above). Secondly, after entering the bacterial cell, and

during in vivo repair, sequences around the site of integration of cDNA into the plasmid are sometimes deleted (see below). Hence some plasmids carry globin sequences which are too small to be detected.

We have therefore carried out a bulk screening of recombinant bacterial clones using a colony filter hybridization technique. As described in detail in Materials and Methods, recombinants were streaked onto nitrocellulose discs on the surface of agar plates, and after overnight incubation the bacterial DNA was immobilized on the filters and hybridized to ^{32}P-labelled globin mRNA. As shown in fig. 6, about 30% of the recombinants tested showed evidence of having incorporated DNA sequences that would hybridize to X. laevis globin mRNA. Some of the positive colonies were then selected for hybridization to globin cDNA in solution.

6. Analysis of recombinants by liquid hybridization to globin cDNA in plasmid excess.

In order to examine the size of the integrated globin DNA sequences, positively scoring recombinant DNAs were analysed for their capacity to protect X. laevis globin cDNA from endonuclease S1 digestion. The results of such an analysis are presented in table 1, the levels of protection rang-

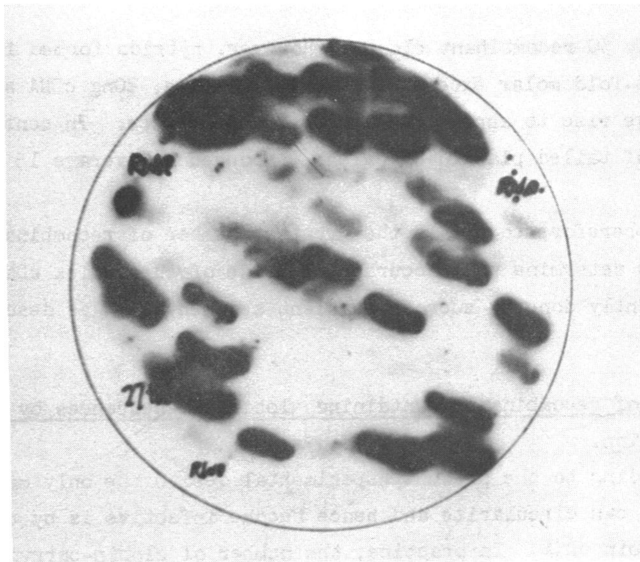

Figure 6 Autoradiograph obtained after hybridization of ^{32}P-labelled globin mRNA to colonies containing recombinant plasmids. The colonies in the top two rows had been scored positive in a previous round of colony hybridization. The other colonies represent a random set of kanamycin resistant transfectants. Exposure, 24 hours.

ing from approximately 5-30% of the input cDNA. Since X. laevis globin mRNA is known to contain two major and at least two minor RNA species (26) the high level of about 30% protection observed with clones B52, C13, C36 and B36 demonstrates that very extensive sequences of X. laevis globin cDNA have been integrated into recombinant plasmids. It was also noted that cDNA hybridization values for the recombinant clones B52 and C13 were additive, suggesting that these recombinants are representative of different species of X. laevis globin mRNA.

7. Heteroduplex mapping of recombinant plasmids

Because the inserted globin sequences were expected to be only a few hundred base pairs, it was anticipated that it would be difficult to identify heteroduplexes between recombinant plasmids and the parental plasmid, pCR1. It was therefore decided to replace pCR1 by a plasmid which would provide a convenient marker for heteroduplex molecules. The plasmid pCR11 is closely related to pCR1 (19) and we have shown that it differs only in the deletion of a 3.9 Kb[a] sequence, located in the ColE1 region, starting 2.1 Kb from the EcoR1 restriction site (unpublished). We have also determined that HindIII cuts both plasmids in the kanamycin resistance gene, 3.5 Kb from the EcoR1 site. From our measurements of the size of pCR11 as 9.2 Kb, we predict that heteroduplexes of HindIII-restricted pCR11 with HindIII-restricted pCR1 recombinant plasmids would show a prominent 3.9 Kb single-stranded loop inserted 39% along the molecule (as measured from the single HindIII restriction target) with the inserted sequences expected to map at the EcoR1 site, 23.1% further along the heteroduplex (i.e. 62.1% from the HindIII site).

TABLE 1

Plasmid designation	% protection of cDNA from S1 digestion	Plasmid designation	% protection of cDNA from S1 digestion
pCR1 B29	21	pCR1 D38	16
B51	13.3	A12	11
B10	6.4	A29	5
C24	12.5	B36	25
B52	30	B52 + C13	50
C13	25	wild-type	0
C36	26		

Solution hybridization of X. laevis globin cDNA to excess recombinant plasmid DNA. For details see text.

Heteroduplexes between clones C13 or B52 and pCR11 (fig. 7) were readily recognised in electron microscope preparations by the presence of the 3.9 Kb deletion loop. Closer examination of such molecules showed a smaller loop at the position predicted for the EcoR1 site, whose size could be

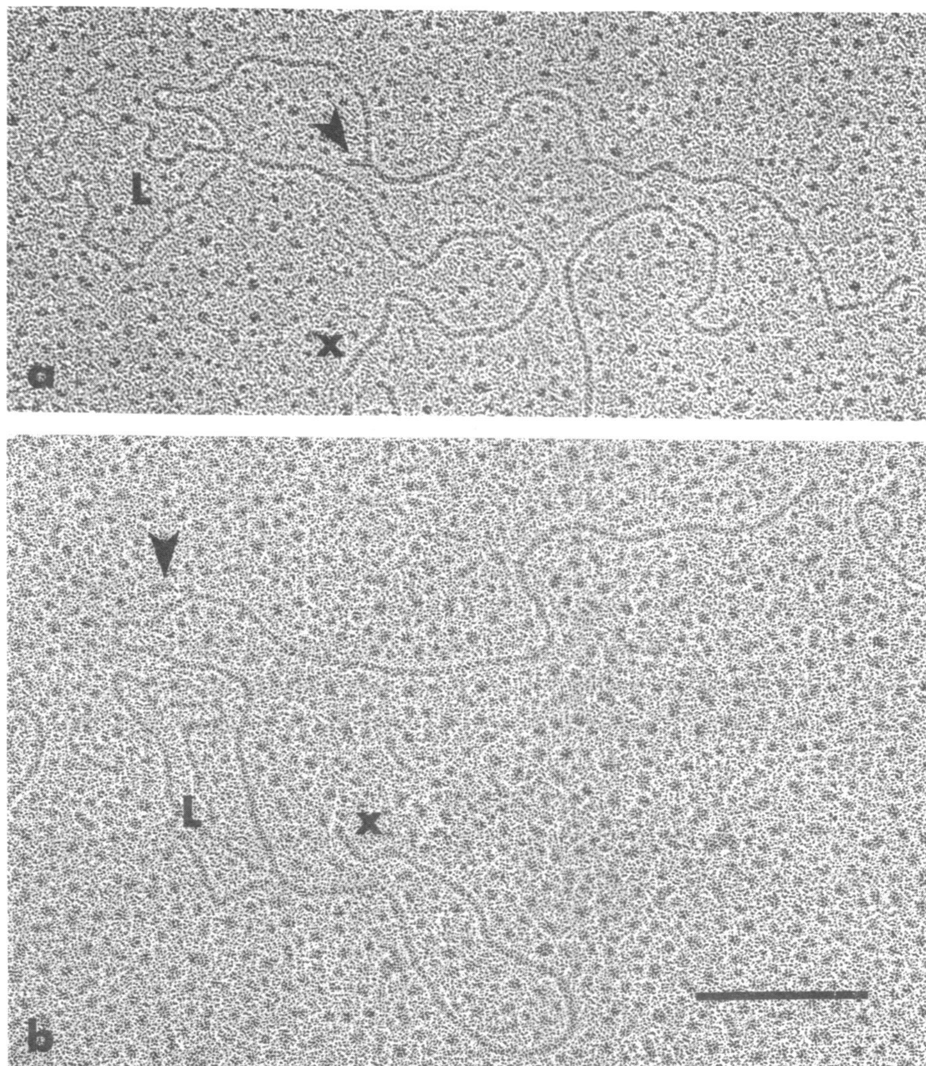

Figure 7 Heteroduplexes of pCR11 with (a) clone B52 and (b) clone C13, all HindIII restricted. Both molecules contain a large single-stranded loop (L) due to the 3.9 Kb deletion in pCR11 about 38% along the molecule from one of the HindIII restricted ends (x). Each heteroduplex also show a smaller loop (arrow), representing the inserted globin sequences and the dG-dC linkers which are absent in pCR11 about 22% further along the molecule. This insertion loop is smaller in the B52-pCR11 heteroduplex than in the C13-pCR11 heteroduplex. The bar represents 0.25μm.

measured using the 3.9 Kb loop as an internal standard. In clone B52 the inserted sequence was determined to be 265$\pm$ 52 base pairs, with the positions of the large and small insertion loops at 37.8$\pm$ 1.2% and 22.8$\pm$ 1.0% respectively (n=20). The inserted sequence in clone C13 was 497$\pm$ 85 base pairs, with the two insertion loops at 38.0$\pm$ 0.65% and 22.5$\pm$ 0.8% respectively (n=20). Neither of the insertion sequences formed an open loop, probably because of their small size and the pairing of the dG-dC linkers used to insert the cDNA sequences.

8. Analysis of recombinant plasmids by polyacrylamide gel electrophoresis.

Fig. 8 illustrates the separation on 3% polyacrylamide gels of

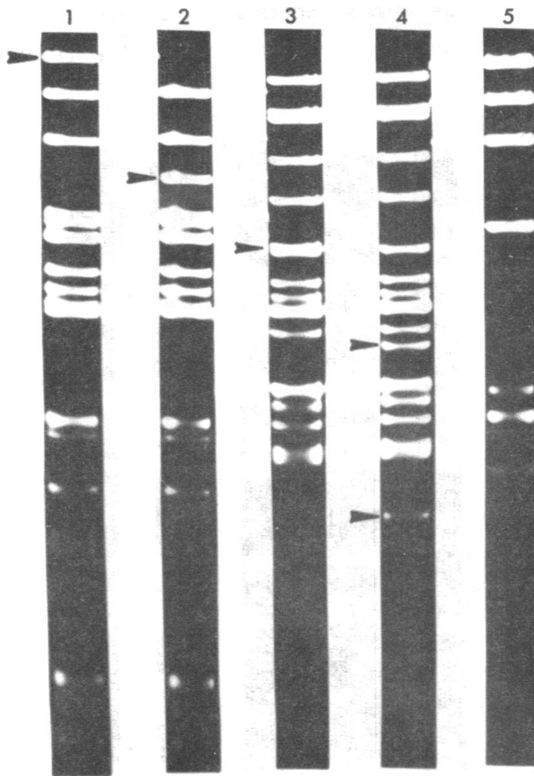

Figure 8 Electrophoresis on 3% acrylamide gels of endonuclease R.Hha and R.HaeIII fragments of pCR1 DNA, with and without further digestion by EcoR1. Gels (0.5 cms diameter) were loaded with 4μg of digested DNA and electrophoresed at 100 volts for 3 hours. The gels were then stained in ethidium bromide and photographed by ultraviolet fluorescence. Lanes 1-4 are Hha, Hha+EcoR1, HaeIII and HaeIII +EcoR1 digests respectively. Lane 5, ϕX174-HaeIII markers. The arrows indicate the changes in mobilities of the fragments due to EcoR1 digestion. The newly-appearing Hha band in lane 2 contains two DNA fragments.

restriction endonuclease Hha and HaeIII fragments of wild-type pCR1 DNA. Both enzymes cleave the DNA into more than 20 fragments. Further digestion of the fragments by EcoR1 resulted in the disappearance of fragments Hha-I and HaeIII-6. This was accompanied by the appearance of new digestion products of lower molecular weight as shown in fig. 8. Hence the incorporation of globin sequences into plasmid DNA can be conveniently studied by observing shifts in the highest molecular weight Hha fragment that contains the single EcoR1 restriction target of the plasmid.

Fig. 9 shows the Hha cleavage patterns of five recombinant plasmids on 3% polyacrylamide gels. In all instances, fragment 1 disappeared and was replaced by fragments of differing mobility. However, other bands also disappeared in some cases. The band affected most frequently was Hha-14 (fig. 9, lanes 2, 5 and 6), although in several recombinants, Hha bands 5 and 7

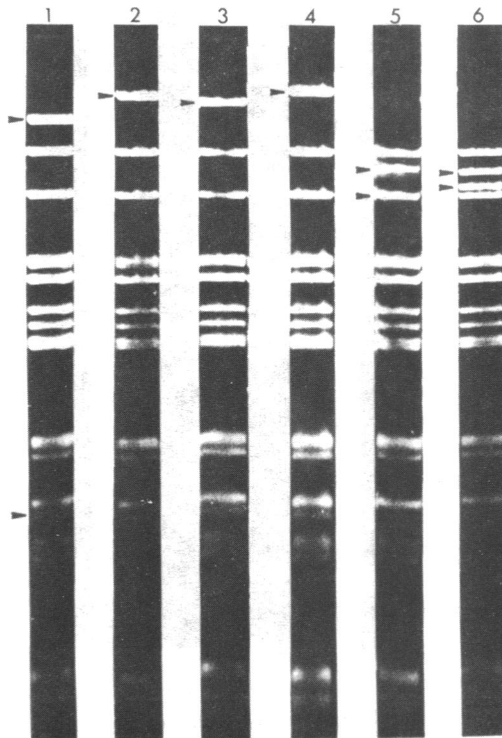

Figure 9 Electrophoresis on 3% acrylamide gels of Hha digests of recombinant plasmids. The gels were run as in figure 8. Lanes 1-6 are wild-type pCR1 DNA and clones B29, C24, B52, B36 and C13 respectively. The arrows indicate the changes in mobilities of fragments due to integration of ds-cDNA into recombinant plasmids.

also disappeared (not shown). Hence, during in vivo repair, extensive DNA sequences were sometimes deleted from recombinant plasmids. Only some of the hybrid plasmids however, showed evidence of being deleted, and it was evident from the shifts in mobilities of the fragments on 3% polyacrylamide gels that recombinants B52, B36 and C13 (fig. 9, lanes 4, 5 and 6) had incorporated in the region of 400 nucleotide pairs of DNA, which is in reasonable agreement with the heteroduplex mapping data (see above).

The recombinant DNAs analysed so far generally fell into two classes, typified by the restriction pattern in fig. 9. While clones B29, C24 and

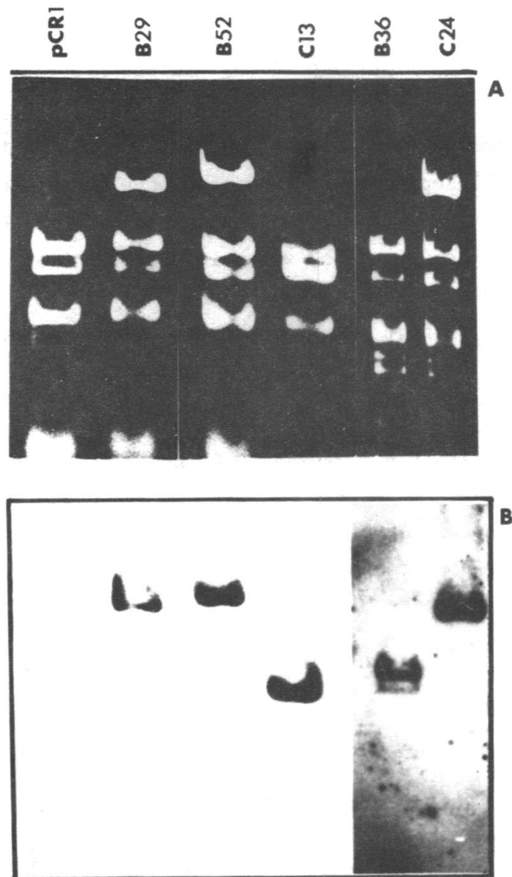

Figure 10 Hybridization of ^{32}P-labelled globin mRNA to DNA fragments in a Hha digest of recombinant plasmids DNAs. About 4μg of plasmid DNA from each indicated clone was digested with Hha, and electrophoresed on 1% agarose gels containing ethidium bromide (0.5μg/ml). DNA bands were denatured, blot-transferred to nitrocellulose filters (27) and hybridized with ^{32}P-labelled globin mRNA. (A) Photograph of ethidium fluorescence. (B) Autoradiographic image obtained after filter hybridization.

B52 (fig. 9, lanes 2, 3 and 4) had one new Hha-1 band of higher molecular weight, clones B36 and C13 (fig. 9, lanes 5 and 6) had two new bands each of smaller size. The second new band in clone B36 (see lower arrow, fig. 9, lane 5) is obscure on this print due to co-migration with the third wild-type Hha band (fig. 9, lane 1). In order to confirm that these newly appearing bands did contain globin-specific sequences, digested plasmid DNAs were electrophoresed in a 1% agarose gel and blot-transferred onto a nitrocellulose filter (27) which was then placed in a hybridization reaction with ^{32}P-globin mRNA. Although the DNA bands were not so well resolved in this gel system as on polyacrylamide gels, the higher molecular weight DNA bands were sufficiently well separated (fig. 10) to show that sequences complementary to globin mRNA were present in the newly-appearing bands of the Hha digests. The highest molecular weight bands of clones B52, B29 and C24 contained a globin sequence while in the case of clone B36 the two new bands were sufficiently well separated to show that both contained globin sequences. The two new bands of clone C13 were not separated in this gel system. One explanation for these observations, which is consistent with the liquid hybridization data (see above), is that clone B52 carries a species of *X. laevis* cDNA different to that carried by clones C13 and B36.

DISCUSSION

In order to facilitate the isolation of genomic fragments of animal DNA it is of considerable advantage to have access to cloned cDNA probes representing the transcribed portions of those genes of interest. The present cloning technique was used because it is the most simple, and the terminal addition of oligo-dGMP to the 3'-OH termini of the ds-cDNA can be monitored with very small quantities of cDNA using the transfection assay described. This could be useful in those instances where only very small amounts of cDNA can be obtained.

It is interesting to note that hybrid plasmids carrying protruding 5'-termini adjacent to the partially single-stranded dG-dC linkages are efficiently repaired after transfection of *E. coli*. The occurance of deleted plasmid sequences in a proportion of recombinant clones is probably not due to the *in vivo* removal of these protruding 5' ends '*per se*', since similar deletions have been observed in recombinant plasmids formed after the *in vitro* repair of *Eco*R1 restricted termini (6). It is possible that some deletions were caused by internal addition of homopolymer to linear plasmids by terminal transferase. The enzyme did contain very low levels of endo-

nucleolytic activity, although no addition of homopolymer to closed circular plasmid DNA could be detected. Alternatively deletions could have arisen during the in vivo repair of hybrid plasmids by the host DNA polymerase.

Globin mRNA from X. laevis is known to consist of two major and at least two minor species all of approximately similar size (26). It is therefore probable that the recombinants obtained from the present experiments are representitive of these four mRNA species. The data suggest that extensive cDNA sequences of two different mRNA species have so far been identified in hybrid plasmids. In order to confirm these observations, individual clones could be assigned to particular globins by using the recently developed technique of hybrid arrested translation (Paterson, B.M., Roberts, B.E. and Kuff, E.L., in the press).

The molecular cloning techniques described in this paper can be used to amplify small quantities of other specific ds-cDNAs, in addition to genomic fragments of animal DNA. It will ultimately be of interest to compare genomic globin sequences in the various species for which specific cloned probes are available.

The molecular cloning experiments described here have been examined by the British Genetic Manipulation Advisory Group and were carried out under Category 2 containment conditions.

[a]ABBREVIATIONS

AMV, avian myeloblastosis virus; MS222, methane tricaine sulphonate; cDNA, single stranded DNA complementary to purified X. laevis mRNA; ds-cDNA, double-stranded cDNA; Kb, kilobase.

ACKNOWLEDGEMENTS

We thank Drs. Mike Houghton, Norman Carey and colleagues of Searle Research Laboratories for a gift of purified endonuclease S1, Dr. F.J. Doos of Medos Company, PortMelbourne, Australia for the provision of foetal calf thymus glands and M. Molloy for typing the manuscript. This work was supported by grants from the Medical Research Council and the Cancer Research Campaign.

REFERENCES

1. Bahl, C.P., Marians, K.J., Wu, R., Stawinsky, J. and Narang, S.A. (1976) Gene 1, 81-92.
2. Seeburg, P.H., Shine, J., Martial, J.A., Baxter, J.D. and Goodman, H.M. (1977) Nature 270, 486-494.
3. Rougeon, F., Kourilsky, P. and Mach, B. (1975) Nucl. Acids Res. 2, 2365-2378.

4. Maniatis, T., Kee, S.G., Efstratiadis, A., and Kafatos, F.C. (1976) Cell 8, 163-182.
5. Rabbits, T.H. (1976) Nature 260, 221-225.
6. Humphries, P., Cochet, M., Krust, A., Gerlinger, P., Kourilsky, P. and Chambon, P. (1977) Nucl. Acids Res. 4, 2389-2406.
7. Sommerville, J. (1977) in International Review of Biochemistry, Biochemistry of Cell Differentiation II, Volume 15, University Park Press, Baltimore.
8. Davidson, E.H. (1977) Gene Activity in Early Development, 2nd Edition, Academic Press, New York and London.
9. Old, R.W., Callan, H.G. and Gross, K.W. (1977) J. Cell Science, 27, 57-79
10. Kacian, D.L. and Spiegelman, S. (1973) Methods Enzymol. 29E, 150-173.
11. Bollum, F.J., Chang, L.M.S., Tsiapalis, C.M. and Dorson, J.W. (1974) Methods Enzymol. 29E, 70-78.
12. Klenow, H., Overgaard-Hansen, K. and Patkar, S.A. (1971) Eur. J. Biochem. 22, 371-381.
13. Green, P.J., Bettack, M.C., Goodman, H.M. and Boyer, H.W. (1974) Methods Mol. Biol. 7, 87-105.
14. Vogt, V.M. (1973) Eur J. Biochem. 33, 192-200.
15. Thomas, N. and Maclean, N. (1975) J. Cell Science 19, 509-520.
16. Harrison, P.R., Birnie, G.D., Hell, A., Humphries, S., Young, B.D., Paul, J. (1974) J. Mol. Biol. 84, 539-554.
17. Roberts, B.E. and Paterson, B.M. (1973) Proc. Nat. Acad. Sci. U.S.A. 70, 2330-2334.
18. Laemmli, U.K. (1970) Nature New Biol. 227, 680-685.
19. Covey, C., Richardson, D. and Carbon, J. (1976) Mol Gen. Genet. 145, 155-158.
20. Grunstein, M. and Hogness, D.S. (1975). Proc. Nat. Acad. Sci. U.S.A. 72, 3961-3965.
21. Denhardt, D. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.
22. Birnie, G.D., MacPhail, E., Young, B.D., Getz, M.J. and Paul, J. (1974) Cell Differentiation, 3, 221-232.
23. Davis, R.W., Simon, M. and Davidson, N. (1971) Methods Enzymol. 21, 413-428.
24. Studier, F.W. (1965) J. Mol. Biol. 11, 373-390.
25. Roychoudhury, R., Jay, E. and Wu, R. (1976) Nucl. Acids Res. 3, 863-877.
26. Battaglia, P. and Melli, M. (1977) Dev. Biology, 60, 337-350.
27. Southern, E. (1975) J. Mol. Biol. 98, 503-517.
28. Richardson, C.C. (1965) Proc. Nat. Acad. Sci. U.S.A., 54, 158-165.
29. Higuchi, R., Paddock, G.V., Wall, R. and Salser, W. (1976) Proc. Nat. Acad. Sci. U.S.A. 73, 3146-3150.