An intron nucleotide sequence variant in a cloned β^+ -thalassaemia globin gene

---

David Westaway* and Robert Williamson

Department of Biochemistry, St. Mary's Hospital Medical School, University of London, London W2 1PG, UK

Received 13 March 1981

ABSTRACT

A 7.5 kb Hsu I restriction fragment of genomic DNA containing a β -globin gene has been isolated from a patient doubly heterozygous for β^+ -thalassaemia and a $\delta\beta$ (Lepore) globin fusion gene. This fragment must be derived from the chromosome carrying the β^+ -thalassaemia determinant. The gross structure of the cloned gene plus flanking sequences is indistinguishable from that of a normal β -globin gene. Within the 1606 base-pair transcribed region of the gene there is only one nucleotide difference from the normal β -globin gene sequence. This is a G→A replacement 21 nucleotides upstream from the 3' terminus of the small intron. This nucleotide lies within a 10 base-pair sequence repeated in an inverted configuration near the 5' terminus of the small intron. The nucleotide replacement may result in a precursor mRNA less amenable to RNA splicing than its normal counterpart.

INTRODUCTION

Thalassaemia is a monogenic recessive hereditary disease common in the Mediterranean, North Africa, and Asia. The disease is characterised by an imbalance in the synthesis of the α - and β -globin chains of adult haemoglobin (HbA; $\alpha_2\beta_2$). In β -thalassaemia there is a deficiency of globin chains, and the disease can be divided into two types, β^0 and β^+ thalassaemia. There are no β chains present in the erythrocytes of homozygous β^0 -thalassaemics, whereas β -globin is present at low levels in homozygous β^+ -thalassaemics (1-3). This β -globin is structurally normal. Levels of β -globin mRNA in homozygous β^+ thalassaemia show a roughly linear correlation with the levels of β -globin protein, and β -globin mRNA isolated from the reticulocytes of these patients is translated at a comparable rate to β -globin mRNA from normal subjects (4-6). This suggests that the disease is the result of a quantitative deficiency in β -globin mRNA. This could be due to mRNA instability. Alternatively the primary molecular defect could be at the level of transcript synthesis or maturation.

Like most eukaryotic genes, the protein coding sequences of the human

globin genes are interrupted by sequences not present in the mature mRNA (7, 8). These tracts of DNA are called "introns" or "intervening sequences". The human β -globin gene contains a small intron (130 base-pairs) between codons 30/31 and a large intron (850 base-pairs) between codons 104/105 (9). The human β -globin gene sequences plus introns are transcribed to give a co-linear precursor mRNA (pre-mRNA) about 1800-2000 nucleotides long (10, 11). The intron sequences are removed from the pre-mRNA by excision/ligation reactions referred to as splicing (12). The splicing, or "processing", of pre-mRNAs occurs in the cell nucleus.

Nienhuis et al (13) have shown by cDNA titration experiments that the steady-state α/β globin mRNA sequence ratio in three homozygous β^+ -thalassaemics was more nearly normal in the nuclei of bone marrow cells than in the cytoplasm or in reticulocyte RNA. Pulse-chase experiments on a total of five β^+ -thalassaemic patients have produced a similar result (10,11). These experiments imply that the β -globin genes are transcribed efficiently in the bone-marrow cells of these patients, but that maturation of nuclear pre-mRNA species is perturbed.

Restriction enzyme mapping has shown that the gross structure of the $\delta\beta$ globin gene locus is unaltered in most cases of β^0 and β^+ thalassaemia (14,15). The disease may be caused by a point mutation, and there is some tentative genetic data that these mutations map near to the β -globin structural gene (1). Maquat et al have suggested that mutations within the β -globin gene introns could produce the abnormal pre-mRNA metabolism observed in some β^+ -thalassaemic patients (10). However, there is also evidence from rarer deletion thalassaemia syndromes and HPFH that distal sequences can affect the expression of the γ - and β -globin genes (16,17). Point mutations within these distal sequences cannot be excluded as possible causes of the common form of β^0 - and β^+ -thalassaemia.

As a first step in identifying the mutation conferring a β^+ -thalassaemic phenotype, a β -globin gene has been isolated from a β^+ -thalassaemic patient. As discussed below, this clone containing the β -globin gene is unequivocally derived from a chromosome which carries the determinant for β^+ thalassaemia. The complete nucleotide sequence of this gene was determined.

MATERIALS AND METHODS

The Patient.

The patient is a 19 year old male of Turkish Cypriot origin who has severe thalassaemia intermedia and presented with a haemoglobin level of 7.2 g/dl. He is transfused regularly. Blood was taken immediately prior to transfusion to ensure that donor white cells did not contribute DNA. In biosynthesis studies the β/α ratio is 0.048 and γ/α is 0.20. The patient is diagnosed as being doubly heterozygous for β^+ thalassaemia and Hb Lepore (18). Hb Lepore was demonstrated in his father by starch gel electrophoresis.

Bacterial Strains.

All phage were grown in the *E.coli* host LE392, a gift from Dr. P. Leder. The Hind III replacement vector NEM788 is a Wam Eam Sam derivative of the phage NEM760 and was a gift from Dr. Noreen Murray (19). The in vitro packaging lysogens BHB2671 and BHB2673 were supplied by Dr. Binie Klein, University of Edinburgh (20). Recombinant phage and subclones containing normal β -globin genes were a gift from Dr. Tom Maniatis and co-workers. Plasmids were grown in *E.coli* HB101 (21). pAT153 was provided by Professor David Sherratt, University of Glasgow (22). Recombinant strains were propagated as advised by the UK Genetic Manipulation Advisory Group.

Construction of Recombinant Bacteriophage.

High molecular weight DNA from the peripheral blood of the patient was prepared as described previously (23). Forty-five micrograms of DNA was digested to completion with the Hind III isoschizomer, Hsu I. The DNA was extracted with phenol, precipitated with ethanol, dissolved in 10mM Tris-HCl pH7.5 1mM EDTA buffer, and electrophoresed on a preparative agarose gel. DNA migrating in the size class between 6.0 and 9.5 kb was located by ethidium bromide staining of size-markers run in parallel. DNA was eluted from the agarose by the "freeze and squeeze" method of Thuring et al (24). λ NEM788 vector DNA was digested with Hsu I, and the central restriction fragment was removed by sucrose gradient centrifugation (25). Size-fractionated human DNA and purified phage vector "arms" were mixed at a molar ratio of 5:1, and ligated at 22°C for 3 h. Ligations were performed in 20mM Tris-HCl pH7.5, 10mM MgCl₂, 20mM 2-mercaptoethanol, 0.5mM ATP, 100 µg/ml enzyme grade bovine serum albumin (Bethesda Research Labs, Inc., Rockville MD). The concentrations of DNA and T4 DNA ligase were 200 µg/ml and 35 Weiss units/ml respectively. The ligated samples were added directly to in vitro packaging aliquots. These aliquots were prepared using the method described by Collins and Hohn (19). Packaging reactions were performed at DNA and ATP concentrations of 20 µg/ml and 6mM respectively. The packaging efficiency was 5×10^5 plaque-forming units per µg of insert DNA. Recombinant phage were plated on NUNC bio-assay dishes without further amplification (26). 1 µl of a low-titre stock of the phage λ H β G2, containing the human δ - and β -globin gene sequences (7), was spotted at two positions on each plate as a control marker. This volume was equivalent to 5-10 phage. The plates were then incubated overnight, chilled, and blotted onto nitrocellulose filters (27). Duplicate filters from each plate were hybridised to a nick-translated genomic Pst 1 fragment excised from a subclone of λ H β G2. This fragment contains a human β -globin gene. Hybridisation and autoradiography were carried out as described previously (23). The 7.5 kb Hsu I and 4.4 kb Pst I fragments of λ 788 β^+ (this paper), were subcloned into the disabled plasmid vector pAT153 using standard methodology.

Enzymes.

Eco RI, Hinf I, and bacterial alkaline phosphatase were from BRL. Hsu

I, Xba I, and Bgl II were prepared by Dr. Janet Arrand (St Mary's Hospital Medical School) and co-workers. All other restriction enzymes were from New England Biolabs, Inc., Beverly, Mass. T4 polynucleotide kinase and T4 DNA ligase were from PL-Biochemicals, Milwaukee, Wisconsin. γ^{32} P-ATP, >2000 Ci/mmol was obtained from the Radiochemical Centre, Amersham, England.

DNA Sequence Analysis.

The chemical modification method of Maxam and Gilbert was used (28,29). The 4.4 kb Pst I subclone of the cloned β^+ -thalassaemia gene, 4.4 β^+ , was used for sequencing. Restriction fragments were dephosphorylated and then labelled at their 5' termini with polynucleotide kinase and γ^{32} P-ATP. Fragments were strand-separated by denaturation and acrylamide gel electrophoresis, and were visualised by autoradiography. Elution from the gel matrix was as described in (29). The ethanol-precipitated DNA was resuspended in water and spun for 30 sec in a microfuge (Eppendorf 5412) to remove any remaining acrylamide fragments. The supernatant was reprecipitated with Na acetate and ethanol, washed with 70% ethanol, and subjected to the G, G+A, C+T, and C specific reactions. For some fragments a T-specific reaction was also used (30). Cleavage products were fractionated on 400mm x 200mm x 0.35 mm acrylamide urea gels run in 75mM Tris-Borate pH8.3, 1.5 mM EDTA buffer. Electrophoresis was at a constant power of 25-30 Joules per second per gel.

RESULTS AND DISCUSSION

It is difficult to distinguish a homozygous β^+ thalassaemic from a double heterozygote for β^+ and β^0 thalassaemia. This ambiguity would not be resolved by molecular cloning alone as both β^+ - and β^0 -thalassaemic β -globin genes are usually superficially indistinguishable from normal β -globin genes (14). For this reason the patient chosen for analysis was a Turkish Cypriot doubly heterozygous for β^+ thalassaemia and the Hb Lepore globin fusion gene (22). The Hb Lepore gene generates different restriction fragments from the β -globin gene (Figure 1). The only β -globin gene that can be cloned using our procedure from this patient's DNA is the one from the chromosome carrying the β^+ -thalassaemic determinant.

Prior to cloning, genomic DNA from the patient was examined by the Southern transfer technique (31). DNA derived from the placenta of a haematologically normal subject was analysed in parallel. The sizes of restriction fragments which hybridise to the cDNA plasmid pH β G1 (32) are summarised in Table 1. This plasmid hybridises to both β - and δ -globin gene sequences. The patient is heterozygous for a 2.6 kb Pst I fragment and a 3.8 kb Xba I fragment (Table 1). These sizes agree closely with previous estimates for fragments derived from Hb Lepore DNA (23). The patient does not appear to be heterozygous for an Hsu I fragment. This is because Hsu I digestion of the Hb Lepore chromosome generates a $\delta\beta$ -globin fragment the size of which is nearly identical to that of the authentic δ -globin fragment (23). These results are consistent with the haematological diagnosis of the patient's phenotype.

Figure 1: Structure of the Patient's β -Globin Loci. chr. = chromosome. Sizes of Pst I fragments are shown in kb. The β -globin gene on the chromosome carrying the β^+ -thalassaemia determinant is labelled β^+ , and the Hb Lepore fusion gene is labelled $\delta\beta$.

The 7.5 kb Hsu I fragment containing the patient's β -globin gene was cloned in the phage lambda replacement vector NEM 788 (18). DNA from the patient was digested to completion with Hsu I. A size-fraction from 6.5 to 9.5 kb was isolated by preparative agarose gel electrophoresis. This fraction excludes the $\delta\beta$ - and δ -globin gene fragments. This DNA was ligated to the purified "arms" of the phage vector and packaged in vitro. Recombinant phage were plated out on 23.5 cm square Petri-dishes. Two spot-titres of the phage λ HG2 were included on these plates. These phage have an inserted fragment containing the linked δ - and β -globin genes. They serve as an internal control in the screening process and can also be used as markers to align duplicate filters blotted from the same plate. 160,000 recombinant phage were screened using a nick-translated genomic β -globin gene fragment as a hybridisation probe. One positive-scoring phage was detected (Figure 2). This phage, designated λ 788 β^+ , was plaque-purified and the inserted 7.5 kb Hsu I fragment was subcloned into the

Table 1: Sizes of globin gene restriction fragments detected in a non-thalassaemic subject and the thalassaemic patient

	Pst 1	Hsu 1	Xba 1
N	4.4 2.3	7.5 18.0	11.0
T	4.4 2.3 +2.6	7.5 18.0	11.0 +3.8

Sizes are given in kb. Southern transfers were performed as described in (23). The hybridisation probe was a β -globin cDNA plasmid, p β 31 (32). N = Normal subject, T = the doubly heterozygous patient.

Figure 2: Screening Recombinant Phage. 1 and 2 are duplicate nitrocellulose filters blotted from one half of a 23.5 x 23.5 cm NUNC Bio-assay dish. This area of the plate contains approximately 25,000 phage. The spot-titre of the control recombinant phage $\lambda H\beta G2$ is circled. The positive-scoring phage, designated $\lambda 788\beta^+$ is arrowed.

plasmid vector pAT153 (21).

The subclone containing the Hsu I fragment, $7.5\beta^+$, was digested with a number of restriction enzymes to determine the physical map shown in figure 3A. The inserted fragment contains a β -globin gene plus approximately 3 kb of 5'- and 3'-flanking sequences. The map of this Hsu I fragment differs from published maps of the normal β -globin gene in only two respects: one extra Pst I site and one extra Bgl II site are present to the 3' side of the gene (7, 16). These "extra" restriction sites are present in subclones of the normal gene, and must have been overlooked in previous analyses. Within the limits of these mapping experiments, about ± 50 base-pairs, this case of β^+ thalassaemia is not associated with the deletion or insertion of DNA sequences in or around the β -globin locus.

The entire β -globin gene was sequenced using the Maxam and Gilbert technique (29, 30). The sequence determined is 1971 nucleotides long and extends 155 nucleotides beyond the "capping" site (34) and 210 nucleotides beyond the poly(A) attachment site. 87% of the sequence has been determined at least twice, and 70% of the sequence has been determined on both strands of the DNA. With the exception of the Eco RI site within the gene, all of the restriction sites used for sequencing have been overlapped (Figure 3B). In addition the availability of a prototype sequence from the normal β -globin gene (9) for cross-checking means that this thalassaemic gene sequence should be highly accurate. Two nucleotide differences from the normal gene sequence have been located (Figure 3C). The first sequence variant lies near the 3' terminus of the small intron. A G residue is replaced by an A residue in the thalassaemic sequence. Both strands of this area of the gene have been sequenced twice, and an identical base-change has been reported in the sequence of an Eco RI β -globin gene fragment isolated from a Greek Cypriot homozygous for β^+ thalassaemia (35).

The $G \rightarrow A$ replacement is not seen in a β -globin gene isolated from a patient doubly heterozygous for $\delta\beta^0$ and β^0 thalassaemia (N. Moschonas and E. de Boer, personal communication). These data confirm that the intron sequence variant is real and is not due to an artefact in the cloning or sequencing of the normal or thalassaemia genes. The second sequence difference is the insertion of an A residue 88 nucleotides beyond the polyadenylation site. Neither of these sequence changes lie within the recognition sequences of any known restriction enzymes, nor do they generate new recognition sequences.

Can the nucleotide sequence of this globin gene be related to the thalassaemic phenotype? The inserted A residue 88 nucleotides beyond the polyadenylation site does not lie in an expressed gene sequence, nor does it map within any of the repetitive elements lying to the 3' side of the β -globin gene (36). It is not obvious how this sequence variant could produce a β^0 -thalassaemic phenotype. The gene codes for a normal β -globin mRNA. Therefore defective mRNA translation can be excluded as the cause of this thalassaemia. Similarly, the 5'- and 3'-flanking sequences, extending for 114 and 88 nucleotides beyond the gene, are identical to the normal sequence. This makes it unlikely that initiation or termination of transcription are perturbed in the thalassaemia gene, although a "long-range" effect on these processes cannot be excluded (17). This gene has not been transcribed *in vitro*, but the 5' Eco RI fragment isolated by Spritz *et al.* is transcribed efficiently *in vitro* (35). The latter fragment has identical 5'-flanking sequences to the gene described here. The homology extends from the Eco RI site at codons 120-121 to at least 155 nucleotides beyond the cap site.

A remaining aspect of gene expression that could be affected in the gene described here is splicing or transport of the pre-mRNA. Splicing of pre-mRNAs has not been investigated in this patient, but has been shown to be anomalous in other β^0 -thalassaemics. The $G \rightarrow A$ replacement is a good candidate for a mutation which could affect these processes. The variant nucleotide lies within an intron, transcripts of which are spliced out from the pre-mRNA. Unfortunately experimental data on the mechanism of intron excision is insufficient to predict whether or not this particular $G \rightarrow A$ replacement could cause ineffective pre-mRNA processing. In two studies insertions or deletions made in intron sequences had no apparent effect on gene transcription and processing to RNA, implying that some intron sequences are functionally silent (37, 38). However internal splice acceptor sites are known to be located within introns, and the $G \rightarrow A$ replacement may alter the activity of such a site (10, 39). The G residue lies within a 10 base-pair sequence which is repeated in an inverted configuration 33-42 base-pairs downstream from the 5' terminus of the small intron (9, 33). Transcription of this inverted repeat sequence will produce a self-complementary RNA molecule which could base-pair to give a stem-loop structure. This type of structure may stabilise intermediates in the splicing reactions.

Figure 3: Structural Analysis of the cloned β -Globin Gene.

A. A restriction map of the insert in the Hsu I subclone 7.5 β^+ . This map was compiled from a total of 24 single and double restriction enzyme digests of the subclone DNA. Electrophoresis on 1.4% agarose gels or 5% acrylamide gels was as described (23). Size-markers were λ DNA restricted with Hsu I and EcoRI, SV40 DNA restricted with Hpa I, and ϕ X174 DNA restricted with Hae III. Coding sequences are indicated by shaded blocks, introns by open blocks. Sequences coding for the untranslated region of the β -globin mRNA are indicated by diagonal shading.

B. Protocol for Sequencing the cloned β -Globin Gene. All sequencing was performed on the PstI subclone, 4.4 β^+ . Three starting fragments isolated by preparative acrylamide gel electrophoresis were used for sequencing (29). These were a

1.9 kb Bam HI fragment containing the gene 5' region, a 0.9 kb Bam HI/Eco RI fragment spanning from codon 99 to codon 121 and including the large intron, and a 1.5 kb Eco RI fragment containing the gene 3' region. The 1.9 kb fragment was digested with Hinf I, or Hph I, or Hae III, kinased labelled and strand-separated. Similarly, the 0.9 kb fragment was digested with Rsa I, or Mbo II, or Mnl I, and the 1.5 kb fragment was digested with Hinf I, or Hph I prior to labelling of the 5' termini. The appropriately sized strand-separated fragments were identified using the known restriction map of the globin gene (9), and were eluted from the matrix of the preparative acrylamide gel (29). A fourth fragment was isolated from a total digest of 4.4 β^+ . This is a 0.19 kb Ava II fragment which spans the junction of the 1.9 and 0.9 kb fragments. Only restriction sites used for sequencing are indicated. Arrows represent the distance sequenced from each restriction site. The blunt end of the arrow is at the labelled 5' terminus. Nucleotides adjacent to the 5' terminus which were not sequenced are indicated by dashed lines.

C. Differences from the Nucleotide Sequence of the Normal β -Globin Gene

Two errors in the normal β -globin gene sequence (9) have been taken into account. These are a T residue instead of an A, and a C residue instead of an A at 83 and 148 nucleotides respectively beyond the polyadenylation site (confirmed by A. Efstratiadis, personal communication). The gene sequences are represented as in 3A. The map positions of the base-changes are indicated by stars above the gene. The relevant nucleotides are shown below the gene, with the normal and thalassaemic sequence on the lower and upper lines respectively. Coding sequences are shown in uppercase letters. The intron/coding block junction was assigned using the GT..AG rule (33).

Conversely the affected nucleotide may lie within an area of the small intron the structure of which is not important for normal processing. Nucleotide replacements in such regions could nonetheless perturb gene expression if they generate novel biologically active sites. Thus Spritz et al. suggest that the G→A replacement creates a new splice acceptor site within the body of the small intron (35). The AG dinucleotide created by the sequence variant is a conserved feature in splice acceptor sites, and the sequence flanking the dinucleotide, TTAGTCT, closely resembles the sequence TTAGGCT at the 3' terminus of the small intron (33, Fig. 3C). The proposed novel acceptor site could thus compete with the authentic acceptor site 20 nucleotides downstream, and consequently retard pre-mRNA processing (35). This possibility could be tested by comparing the splicing of the normal and thalassaemic gene products in a functional assay. If the G→A base replacement is responsible for anomalous splicing activity, then normal splicing should be recovered on reverting the affected A to G by site-directed mutagenesis. Alternatively, the demonstration of the G→A replacement in a clinically normal subject would establish that the

replacement is an asymptomatic sequence polymorphism.

It has been suggested that sequencing of normal and thalassaemic globin genes would also reveal DNA sequence variants unconnected with the anaemia, and that this genetic "noise" would make detection of the primary lesion problematic (3, 40). The sequence of this, and other β -globin genes isolated from thalassaemic patients (35, N. Moschonas and E. de Boer, personal communication) demonstrates that these genes are highly conserved between unrelated individuals. Only two variable bases have been found within 1971 base-pairs of DNA sequenced here. A previous estimate that 1 in 100 base-pairs in the human genome will vary polymorphically may be an overestimate for the β -globin gene, but may still be generally applicable (40). Only one of the sequence variants identified is a reasonable candidate for the primary lesion in this genetic disease. Further functional studies are needed to assess the importance of this sequence variant, and these are in progress.

ACKNOWLEDGEMENTS

We particularly thank Dr. B. Forget, and Drs. N. Moschonas, E. de Boer and R. Flavell for communicating and discussing gene sequences prior to publication, Peter Little and Ian Jackson for many useful discussions, and Bernadette Modell for pointing out the compound heterozygote in our stocks of human DNA. This work was supported by grants from the British Medical Research Council and the National Institutes of Health (1R01AM20125-01A1).

*Present address: Department of Microbiology and Immunology, School of Medicine, University of California at San Francisco, San Francisco, CA 94143, USA. Reprint requests to this address.

ABBREVIATIONS: Hb = Haemoglobin, kb = kilobases.

REFERENCES

1. Weatherall, D.J. and Clegg, J.B. (1972) The Thalassaemia Syndromes (Blackwell, Oxford, 2nd Ed.).
2. Forget, B.G. (1978) *Trend. Biochem. Sci.* **3**, 86-89.
3. Bank, A., Mears, J.G. and Ramirez, F. (1980) *Science* **207**, 486-493.
4. Nienhuis, A.W. and Anderson, W.F. (1971) *J. Clin. Invest.* **50**, 2458-2460.
5. Reider, R.F. (1972) *J. Clin. Invest.* **51**, 364-372.
6. Benz, E.J. Jr., Forget, B.G., Hillman, D.G., Cohen-Solal, Pritchard, J., Cavallero, C., Prenskey, W. and Housman, D. (1978) *Cell* **14**, 299-312.
7. Lawn, R.F., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) *Cell* **15**, 1157-1174.
8. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., deRiel, J.K., Forget, B., Weissman, S., Slightom, J.L., Blechl, A.E., Smithies,

- O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) *Cell* 21, 653-668.
9. Lawn, R.M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980) *Cell* 21, 647-651.
10. Maquat, L.E., Kinniburgh, A.J., Beach, L.R., Honig, G.R., Lazerson, J., Ershler, W.B. and Ross, J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77, 4287-4291.
11. Kantor, J.A., Turner, P.H. and Nienhuis, A.W. (1980) *Cell* 21, 149-157.
12. Chow, L.T., Gelinas, R.E., Broker, T.R. and Roberts, R.J. (1977) *Cell* 12, 1-8.
13. Nienhuis, A.W., Turner, P. and Benz, E.J. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 3960-3964.
14. Flavell, R.A., Bernards, R., Kooter, J.M., De Boer, E., Little, P.F.R., Annison, G. and Williamson, R. (1979) *Nuc. Acids Res.* 6, 2749-2760.
15. Orkin, S.H., Old, J.M., Weatherall, D.J. and Nathan, D.G. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 2400-2404.
16. Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1979) *Nature* 279, 598-603.
17. Van der Ploeg, L.H.T., Konings, A., Oort, M., Roos, D., Bernini, L. and Flavell, R.A. (1980) *Nature* 283, 637-642.
18. Murray, N.E., Brammar, W.J. and Murray, K. (1977) *Molec. gen. Genet.* 150, 53-61.
19. Collins, J. and Hohn, B. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 4242-4246.
20. Boyer, H.W. and Roulland-Dussoix, D. (1969) *J. Mol. Biol.* 41, 459-472.
21. Twigg, A.J. and Sherratt, D.J. (1980) *Nature* 283, 216-218.
22. Baglioni, C. (1962) *Proc. Natl. Acad. Sci. U.S.A.* 48, 1880-1884.
23. Flavell, R.A., Kooter, J.M., De Boer, E., Little, P.F.R. and Williamson, R. (1978) *Cell* 15, 25-41.
24. Thuring, R.W.J., Sanders, J.P.M. and Borst, P. (1975) *Anal. Biochem.* 66, 213-220.
25. Maniatis, T., Hardison, R.C., Lacy, E., Lamer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) *Cell* 15, 687-701.
26. Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M. and Tonegawa, S. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 4709-4713.
27. Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180-182.
28. Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 560-564.
29. Maxam, A.M. and Gilbert, W. (1980) *Methods in Enzymology* 65, Part 1, 499-560.
30. Rubin, C.M. and Schmid, C.W. (1980) *Nucleic Acids Res.* 8, 4613-4619.
31. Southern E.M. (1975) *J. Mol. Biol.* 98, 503-517.
32. Little, P., Curtis, P., Coutelle, Ch., Van den Berg, J., Dalgleish, R., Malcolm, S., Courtney, M., Westaway, D. and Williamson, R. (1978) *Nature* 273, 640-643.
33. Breathnach R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 4853-4857.
34. Baralle, F.E. (1977) *Cell* 12, 1085-1095.

35. Spritz, R.A., Jagadeeswaram P., Biro, P.A., Elder, J.T., Gefter, M.L., Weissman, S.M. and Forget, B.G. (1980) Proceedings of the NIH Hemoglobin Switching Meeting, Airlie House, Va., in press.
36. Coggins L., Grindlay, G.J., Vass, J.K., Slater, A.A., Montague, P., Stinson, M.A. and Paul, J. (1980) Nuc. Acids Res. 8, 3319-3333.
37. Johnson, J.D., Ogden, R., Johnson, P., Abelson, J. and Itakura, K. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2564-2568.
38. Volckaert G., Feuteun, J., Crawford, L., Berg, P. and Fiers, W. (1979) J. Virol. 30, 674-682.
39. Kinniburgh, A.J. and Ross, J. (1979) Cell 17, 915-921.
40. Jeffreys, A.J. (1979) Cell 18, 1-10.