Volume 9 Number 17 1981 Nucleic Acids Research

Structure and expression of a cloned β^0thalassaemic globin gene

N.Moschonas, E.de Boer, F.G.Grosveld, H.H.M.Dahl, S.Wright, C.K.Shewmaker and R.A.Flavell

Laboratory of Gene Structure and Expression, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Received 2 July 1981

ABSTRACT

We have cloned the single β-globin gene from an Italian patient who is a double heterozygote for $\beta^0/\delta\beta^0$ thalassaemia. RNA isolated from nucleated red cells from this patient can be translated in vitro to give detectable levels of Aγ- Gγ and α-globin, but no β-globin. S_1-mapping transcription studies show that β-globin mRNA is present at about 1-3 % of the level of α- and γ-globin mRNA. In addition, the expression of this gene has been studied by reversed genetics. SV40-plasmid-β^0-globin gene recombinants have been transfected into Hela cells and analysed for β-globin mRNA. In contrast to the results obtained with mRNA isolated directly from the blood of this patient, in the transfected Hela cells the same level of β-globin mRNA is seen for the β^0 thalassaemic globin gene and for a normal β-globin gene.

To elucidate the nature of the lesion, the entire DNA sequence of the β-globin gene of this patient has been determined. The sequence shows that this gene contains a termination codon at position 39 (CAG -> UAG). Otherwise, there is a remarkable conservation of the entire DNA sequence.

INTRODUCTION

The hereditary disease β-thalassaemia is characterized by a reduced level of β-globin chains; since the α-globin chains are produced at the normal level, an imbalance in the synthesis of the α and β-chains ensues. β-thalassaemia can be divided into two types, β^0thalassaemia, where no β-globin chains are produced, and β^+thalassaemia, where low levels of structurally normal β-globin chains can be found (1-3).

Restriction enzyme mapping data have shown that the gross structure of the δ β-globin gene region is unaltered in most cases of β-thalassaemia (4, 5), although in one form (4, 5) the 3' terminal exon of the β-globin gene is deleted. In another form of β^0thalassaemia, Kan and his colleagues (6) have shown that globin mRNA is present at an approximately seven-fold reduced level, relative to α-globin mRNA and that the mRNA contains a translational stop codon at position 17 (AAG -> UAG) (7). Other forms of β^0-thalassaemia

have been characterized with levels of β-globin mRNA that could not be detected by cDNA titration studies.

To determine the nature of the lesion in this type of β^{o}-thalassaemia, we have cloned the β-globin gene from a patient who has been studied previously (8). We have analyzed the globin mRNA directly from the patient as well as analysing the expression of the β-globin gene using SV40 recombinants. DNA sequence studies show that this patient too, has a translational stop codon, this time at position 39 (CAG -> UAG).

MATERIALS AND METHODS

Cloning of the β-globin gene from a patient with $\beta^{o}/\delta\beta^{o}$ thalassaemia

DNA from the $\beta^{o}/\delta\beta^{o}$ thalassaemia patient was isolated from blood as previously described (19). 10.8 mg of the DNA was then digested to completion with restriction endonuclease Hind III and fractionated by electrophoresis on a 0.5% preparative agarose gel. Fractions were phenol extracted, ethanol precipitated and resuspended in 10 mM Tris/HCl, pH 7.5. Fractions containing the β-globin gene fragment were determined by Southern blot analysis of aliquots of the fractions using nick translated plasmid HβG-DNA as probe (19). Positive hybridization was obtained with DNA from fractions containing DNA in the 7-9 kb size range.

Phage λ Ch21A was grown and DNA isolated as published (e.g. 20). 100 μg λ DNA (in 100 mM Tris/HCl, 10 mM $MgCl_2$, pH 8.0) was annealed at 42°C for 1 hr. followed by ligation at 37°C for 3 hrs in a buffer containing 100 mM Tris/HCl, 10 mM $MgCl_2$, 10 mM DTT, 1mM ATP, 0.02% gelatin and 1000 U T4 DNA ligase. The λ Ch21A DNA was then digested to completion with HindIII, treated with bacterial alkaline phosphatase, phenolextracted, ethanol precipitated and resuspended in 10 mM Tris/HCl, pH 7.5.

Human DNA fragments (containing the β-globin gene) were ligated to the vector DNA in the ratio 1:3, packaged *in vitro* as described by Hohn and Murray (21) and plated on bacterial strain DP50 sup F (20). 200,000 plaques were screened according to the method of Benton and Davis (22). One plaque was found to be positive on rescreening. The human β-globin Hind III DNA fragment was then subcloned in the Hind III site plasmid pBR322.

Detection of globin mRNA by the S_1 nuclease mapping

RNA from blood cells from the patient or from Hela cells was extracted as described (13) and hybridized to DNA probes as described previously (14). The probe for the α-globin gene product was a 5' labelled HindIII fragment (from a Sst I fragment of the α-globin gene subcloned in pBR322)which should

give a 179 nucleotide product corresponding to the fragment from the HindIII site in the second exon to the splice acceptor of the second exon. The probe was labelled as described (14). The γ-globin probe was a 3' labelled EcoRI fragment from the γ globin gene. This fragment was recovered from a cloned γδβ globin gene cosmid (cosHG28; see ref. 15). This probe should generate a 167 NT long S_1 nuclease resistant fragment which extends from the EcoRI site in the 3rd exon to the 3' end of the γ globin gene. The β-globin probe was a 3' labelled EcoRI-MspI double digest fragment labelled at the EcoRI site. Hybridization to β-globin mRNA generates a 212 NT S_1 nuclease resistant fragment. The 3' labelled probes were prepared as described previously (13) using α-^{32}P labelled dATP and TTP(2000/3000 Ci/mMol, Amersham).

Transformation of Hela cells with β-globin gene-SV40 recombinants

A BamHI-EcoRI fragment of SV40 was subcloned in pBR328 (a generous gift of G.C. Grosveld). The 4.7kb BglII fragment, which contains the β-globin gene was isolated from a normal (15) or a β^othalassaemic individual (this report) and subcloned in the BamHI site of the SV40-pBR328 recombinant. 15 μg of the recombinant DNA was applied to Hela cells (grown in α-MEM + 10% newborn calf serum) as a calcium phosphate coprecipitate as described (15). After removal of the excess DNA-precipitate after about 16 h, the cells were grown for 48h, RNA isolated (13) and the level of β-globin mRNA determined by S_1 mapping. Under these conditions the rabbit β-globin gene generates about 10^4 β-globin mRNA molecules per cell.

RESULTS

Expression of the β^othalassaemic globin gene in vivo

Most studies of β^othalassaemia have been performed on "homozygous" patients. Though technically simple, this approach suffers from the disadvantage that if each respective allele exhibits a different form of β^o thalassaemia, then it is not usually possible to distinguish these in in vivo studies of gene expression. The general consensus is that β-thalassaemia is a heterogeneous disease which may be a manifestation of several different types of defect (see e.g. refs. 1-3). For this reason, we chose to study the β^o-globin gene of a patient doubly heterozygous for $\beta^o/\delta\beta^o$ thalassaemia. Since in δβothalassaemia the β-globin gene is entirely deleted (9,10,11),all studies of β-globin gene expression must therefore only refer to the single remaining β-globin gene in this patient (see Fig. 1).

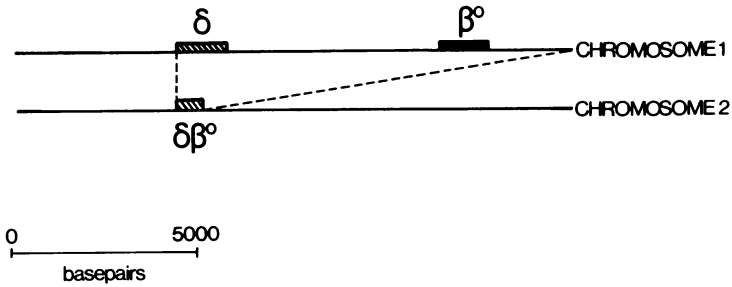

Fig. 1. Schematic genotype of the patients at the δ and β-globin locus. chr. = chromosome. $\delta\beta^{o}$ refers to the partially deleted $\delta\beta^{o}$ thalassaemic globin gene and β^{o} to the thalassaemic β-globin gene. Blotting experiments described previously have shown that this patient has the classical Southern Italian form of $\delta\beta^{o}$-thalassaemia depicted here (9).

We first isolated total RNA from this patient and translated this *in vitro* using a wheat germ extract to establish that no translatable β-globin mRNA is present. The translation product was analysed on Triton-urea gels (12), since these separate both Gγ-, Aγ-, β-and α-globins; while Gγ, Aγ and α-globins were abundantly produced, no β-globin could be detected (not shown).

The RNA preparation was then assayed for α-, β-, γ- and δ-globin mRNA by hybridizing the preparation with 3' end-labelled probes for the β-γ- and δ-globin genes and a 5' probe for the α-globin gene as described in Materials and Methods. The hybrids were treated with S_1 nuclease and the S_1-resistant products analysed on polyacrylamide-urea 'sequencing' gels to determine the size of the product. Previous studies on the globin mRNA of this patient showed that the level of β-globin mRNA was below the detection level of the cDNA titration method used although α-globin mRNA was shown to be present (7). Using S_1-mapping, we found high levels of α- and γ-globin mRNA (Fig. 2) and a low level of S_1 resistant hybrid was formed with a δ-globin gene probe (not shown). Low but significant levels of β-globin mRNA are, however, also detectable in the RNA of this patient with this method. Both the 5' (not shown) and 3' ends (Fig. 2) of this mRNA are the same as those of normal β-globin mRNA, since the S_1 nuclease resistant products have the same size as those obtained with normal human β-globin mRNA. The extensive sequence divergence of the δ- and β-globin genes in the 3' untranslated region precludes the detection of heterologous δ-β hybrids with the β-globin

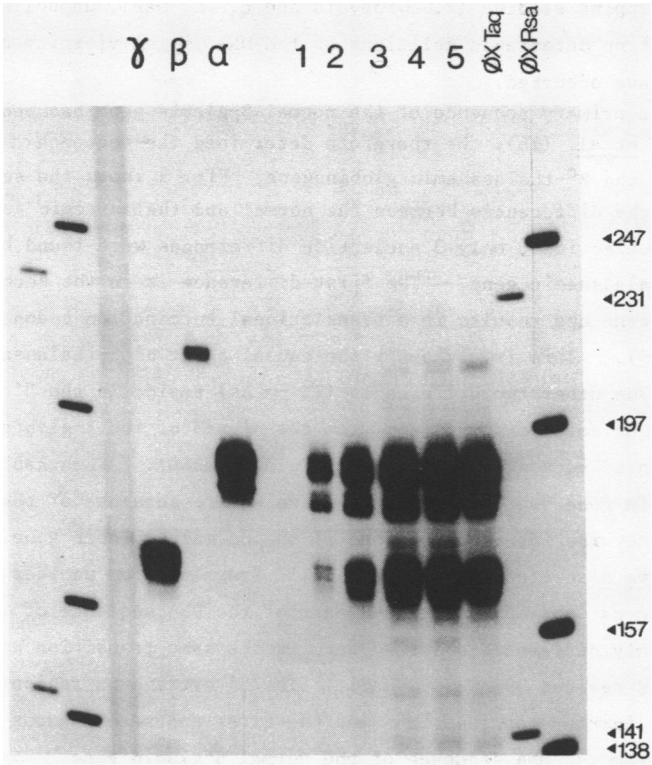

Fig. 2. Determination of the α, β and γ-globin mRNA levels in whole blood RNA from this patient. The RNA levels were determined by hybridization to 3' end labelled probes for the β (2×10^4 cpm/p mol 3' ends) and γ-globin genes (1.72×10^4 cpm/p mol 3' ends) and to a 5' labelled probe for the α-globin gene (1.5×10^4 cpm/p mol 5' ends). On the left side of the figure the probes have been hybridized separately to an excess of total blood RNA in a volume of 10 µl. On the right side the three probes were mixed and hybridized to 4ng, 20ng, 100ng, 400ng and 2µg respectively (1 to 5), of total RNA. The markers are 5' labelled fragments of ØX 174 DNA cut with TaqI or RsaI.

gene probe using this method.

Cloning of the thalassaemic β^o-globin gene and determination of the β-globin gene DNA sequence

A variety of mechanisms could be envisaged to explain the low level of β-globin mRNA and the absence of β-globin in the red cell precursors of this patient. To determine the molecular basis of the lesion in this case, we cloned a 7.5kb HindIII fragment which contains the β-globin gene using as vector bacteriophage λ charon 21A (see Materials and Methods). Standard

restriction mapping studies (F.G.Grosveld and H.H.M. Dahl, unpublished) confirmed that no detectable deletions of the DNA in the vicinity of the β-globin gene have occurred.

The primary sequence of the normal β-globin gene has been determined by Lawn et al. (16). We therefore determined the sequence of the same DNA region of the β^{o}-thalassaemic globin gene. Fig. 3 shows the sequencing strategy and the differences between the normal and thalassaemic gene sequence. Out of 2043 nucleotides, only 3 nucleotide differences were found between the normal and thalassaemic gene. The first difference is in the second exon of the β-globin gene and results in a translational termination codon at residue 39 (CAG -> UAG). This is obviously the causal agent of β-thalassaemia in this case. The other two differences (TT -> AA) reside in the 3' extragenic regions, 214 nucleotides downstream from the 3' end of the β-globin gene (defined as those sequences that encode β-globin mRNA). Remarkably, the entire β-globin gene sequence, including the entire sequence of the large intron, is otherwise identical to that of the normal β-globin gene.

We have also cloned the same HindIII fragment from another Italian β^{o}thalassaemic patient and determined part of the DNA sequence of the β-globin gene. The only difference observed here is the same transition at the codon for amino acid residue 39 (CAG -> UAG). The 5' extragenic region, the 5' and 3' intron-exon junctions and the first and second exons show no differences from the DNA sequence of the normal β-globin gene.

Expression of the β^{o}thalassaemic globin gene on an SV40 recombinant

The facts that the transcribed portion of the β^{o}-thalassaemic globin

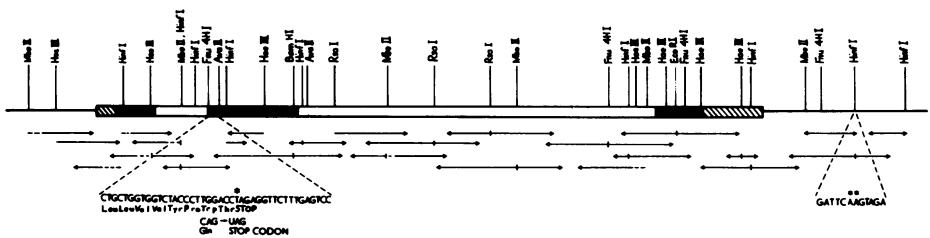

Fig. 3. The sequencing strategy for the β^{o}thalassaemic globin gene and the sequence differences observed. The β-globin gene is indicated as protein coding (filled boxes) 5' or 3' untranslated sequences (hatched boxes) or intron (open boxes). The TT -> AA difference in the 3' extragenic regions may not be real. The sequence of this region of the normal β-globin gene requires reconfirmation (T. Maniatis, A. Efstratiadis and N.Proudfoot, personal communication).

gene differs from the normal gene by only a single base substitution and that the 5' flanking DNA sequences are identical for both genes (see Discussion), suggest that the translational stop codon causes the low level of β-globin mRNA in the thalassaemic. Alternatively, it could be argued that a second mutation exists in the vicinity of the thalassaemic β-globin gene which causes, for example, a reduction in the level of transcription.

To examine this possibility we linked the β^{o}thalassaemic globin gene to an SV40-pBR328 plasmid vector which we have recently used to study the expression of the rabbit β-globin gene (G. C. Grosveld, E. de Boer, C. K. Shewmaker and R. A. Flavell, in preparation). The 4.7kb BglII fragment (which contains the human β-globin gene together with about 2 kb of 5' flanking DNA sequences and about 1.2 kb of 3' flanking DNA sequences), was introduced into the BamHI site of the SV40 vector by cloning in E.coli as described in Materials and Methods. This inserts the β-globin gene segment between the early region of SV40 and the tetracycline gene of pBR328. Both the normal and β^{o}thalassaemic globin genes were inserted into the SV40 vector and plasmids containing the human β-globin genes in both orientations (i.e. with the 5' end of the β-globin gene adjacent to the SV40 and pBR328 sequences respectively) were obtained.

The SV40 β-globin gene recombinants were then introduced into Hela cells as a calcium phosphate co-precipitate (see Materials and Methods). Under these conditions, W. Schaffner (personal communication) has shown that the rabbit β-globin gene inserted into the SV40 KpnI site of an SV40-pBR322 plasmid, is expressed specifically from its own promotor and that translatable β-globin mRNA is produced in these cells. To detect the human β-globin gene transcripts in the Hela cells, a 3' labelled 700 bp EcoRI + MspI fragment, which spans the 3' end of the β-globin gene was used as a hybridization probe in S_1 mapping experiments as described above and in Materials and Methods. Fig. 4 shows that the expected DNA fragment is also found when RNA from the Hela cells that have been transfected with human β-globin gene is hybridized with the 3' labelled fragment and digested with S_1 nuclease. Moreover, approximately the same level of β-globin mRNA is produced from the normal or thalassaemic β-globin gene. No protected fragment is seen for control Hela cells (not shown) or Hela cells transformed with the SV40 pBR328 vectors (Fig. 4). It is therefore unlikely that the low level of β-globin mRNA in the thalassaemic patient is due to a strongly reduced transcription level in vivo.

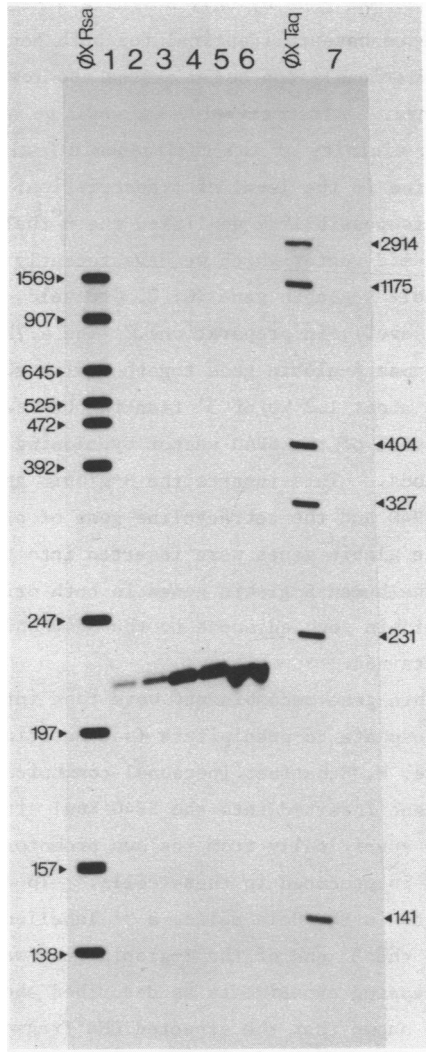

Fig. 4. β-globin mRNA levels produced from either the thalassaemic or normal β-globin gene as SV40 recombinants in Hela cells. The level of β-globin mRNA was determined by S_1 nuclease mapping as described in Materials and Methods using as probe a 3' labelled 0.7kb EcoRI + MspI fragment of the β-globin gene. The hybrids were analysed on an 8O% polyacrylamide urea sequencing gel. As markers ^{32}P 5' labelled ØX DNA x RsaI and ØX x TaqI were used. Hybridization and S_1 nuclease treatment was performed as described in Materials and Methods. The RNA samples hybridized were 5µg 'normal' (1) or 'thalassaemic' (2) 20 µg 'normal' (3) or 'thalassaemic' (4) or 5Oµg of 'normal' (5) or 'thalassaemic' (6). As a control, RNA from Hela cells transformed with the SV4O-pBR328 vector (7) was used. The number refers to the molecular weights of the marker fragments.

DISCUSSION

In this study we have employed a combination of structural and functional analyses to elucidate the nature of the lesion in a case of β^{o} thalassaemia. The DNA sequence of the thalassaemic β-globin gene in these two Italian patients, shows that the primary lesion is a translational termination codon at residue 39; Kan and his colleagues (6) previously showed a termination codon at residue 17 in a Chinese form of β^{o} thalassaemia and more recently, they have demonstrated an identical lesion to the one described here (CAG -> UAG) in another Italian β^{o} thalassaemic patient (25).

It seems possible, therefore, that this will turn out to be a common lesion in β^{o} thalassaemia. It should be noted that while earlier studies suggested that these patients did not produce β-globin mRNA (e.g. 11), these current studies show this RNA to be present. Whether β-thalassaemic cases with zero levels of β-globin mRNA exist remains to be seen.

Interestingly, in the patient we have characterized in most detail, the level of β-globin mRNA is only about 1-3% of the normal level· it follows that β-globin mRNA level is about 50 fold lower than the normal value. Y. W. Kan and his colleagues have made a similar observation for their patient (25). Since the entire β-globin gene only differs at one nucleotide position, it would follow that the β-globin mRNA produced only differs at a single residue, namely in the translational stop codon and that this is the cause of the low level of β-globin mRNA. The alternative explanation is, of course, that an additional mutation has occurred, for example, in the β-globin gene promotor which results in a reduced rate of RNA synthesis. We do not consider this likely. The 5' extragenic nucleotide sequence of the β^{o} thalassaemic globin gene is identical to that of the normal β-globin gene up to residue -340 (this report and N.M., E.de B. and R.A.F., unpublished data). This is beyond the region considered to contain the β-globin gene promotor (for discussion, see 18). More important, the β-thalassaemic gene generates the same level of β-globin mRNA as is found for the normal β-globin gene, when these genes are transcribed in Hela cells.

At this stage it seems likely that the low level of β-globin mRNA in the erythroblasts of this patient may be a result of a reduced stability of this defective β-globin mRNA, as was previously proposed for the Chinese β^{o} thalassaemia (6). Such instability might result from the nuclease susceptibility of mRNAs which are not protected by ribosomes because protein synthesis terminated prematurely on these mutant mRNAs. Since this difference

is not seen in relatively short-term (48 h) in vivo experiments in Hela cells, we believe that it results from a difference between the Hela cell and in vivo nucleated red cell cytoplasm that must at present remain the subject of speculation.

Finally, it is noteworthy that exceedingly few polymorphic differences exist between the DNA sequence of this β^{o}thalassaemic β-globin gene and the normal β-globin gene sequence. In fact, of the total of 1600 nucleotides present in the transcribed segments of each of four β-globin genes (one β^{o} thalassaemic (this report), two identical β^{+}thalassaemics (23,24)and one normal gene (17)), only one nucleotide difference is seen per abnormal gene and this difference appears to be the causal agent of the disease in each case - no polymorphic differences have been seen. This differs markedly from the human foetal globin genes where two allelic Aγ globin genes differed by 15 base substitutions and three deletions of 18, 4 and 4 nucleotides respectively.

ACKNOWLEDGEMENTS

F.G.G., H.H.M.D., E.deB. and R.A.F. are grateful to J. Collins for his hospitality in Braunschweig in the earliest stages of this work, to T. Maniatis for a gift of the α-globin genomic subclone, and to R. Harvey and A. Smith for help with the in vitro translation of globin mRNA. F.G.G. and R.A.F. gratefully acknowledge an EMBO short term fellowship to fund the original visit to Braunschweig. H.H.M.D. and N.M. were recipients of EMBO long term fellowships. F.G.G. and C.K.S. were recipients of a Royal Society European Programme fellowship and a Jane Coffin Childs fellowship respectively. This work was supported by the British Medical Research Council.

REFERENCES

1. Weatherall, D. J. and Clegg, J. B. (1972) in: The Thalassaemis Syndromes. Blackwell Scientific Publications, Oxford (second edition).
2. Forget, B. G. (1978). Trends Biochem. Sci. 3, 86-89.
3. Bank, A., Mears, J. G. and Ramirez, F. (1980). Science 207, 486-493.
4. Flavell, R. A., Bernards, R., Kooter, J. M., de Boer, E., Little, P.F.R., Annison, G. and Williamson, R. (1979). Nucl. Acids Res. 6, 2749-2460.
5. Orkin, S. H., Old, J. M., Weatherall, D. J. and Nathan, D. G. (1979). Proc. Natl. Acad. Sci. U.S.A. 76, 2400-2404.
6. Temple, G. F., Chang, J. C. and Kan, Y. W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3047-3051.
7. Chang, J. C. and Kan, Y. W. (1979). Proc. Natl. Acad. Sci. U.S.A. 76, 2886-2889.

8. Ottolenghi, S., Lanyon, W. G., Williamson, R., Weatherall, D. J. Clegg, J. B. and Pitcher, C. S. (1975). Proc. Natl. Acad. Sci. U.S.A., 72, 2294-2299.
9. Bernards, R., Kooter, J. M. and Flavell, R. A. (1979). Gene 6, 265-280.
10. Fritsch, E. F., Lawn, R. M. and Maniatis, T. (1979) Nature, 279, 598-603.
11. Ottolenghi, S., Giglioni, B., Comi, P., Gianni, A. M., Polli, E., Acquaye, C. T. A., Oldham, J. H. and Masera, G. (1979) Nature, 278, 654-657.
12. Alter, B. P., Goff, S. C., Efremor, G. D., Gravely, M. A. and Huisman, T. H. (1980). Brit. J. Haematol. 44, 527-534.
13. Auffrey, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314..
14. Grosveld, G. C., Koster, A. and Flavell, R. A. (1981) Cell 23, 573-584.
15. Grosveld, F. G., Dahl, H. H. M., de Boer, E. and Flavell, R. A. (1981) Gene 13, 227-237.
16. Wigler, M., Silverstein, S., Lee, L-S., Pellicer, A., Cheng, Y. and Axel, R. (1977) Cell 11, 223-232.
17. Lawn, R. M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980) Cell 21, 647-651.
18. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. and Proudfoot, N. J. (1980) Cell 21, 653-668.
19. Flavell, R. A., Kooter, J. M., deBoer, E., Little, P. F. R. and Williamson, R. (1978) Cell 15, 25-41.
20. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. and Efstratiadis, A. (1978) Cell 15, 687-701.
21. Hohn, B. and Murray, K. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3259-3263.
22. Benton, W. and Davis, R. (1977) Science 196, 180-182.
23. Westaway, D. and Williamson, R. (1981) Nucl. Acids Res., 9, 1777-1788.
24. Spritz, R. A., Jagadeeswaran, P., Choudary, P. V., Biro, P. A., Elder, J. T., deRiel, J. K., Manley, J. L., Gefter, M. L., Forget, B.G. and Weissman, S. M. (1981) Proc. Natl. Acad. Sci. U.S.A. (in press).
25. Trecartin, R. F., Liebhaber, S. A., Chang, J. C., Lee, K. Y., Kan, Y. W., Furbetta, M., Angius, A. and Cao, A. (1981). J. Clin. Invest. (in press).