Proc. Natl. Acad. Sci. USA
Vol. 78, No. 7, pp. 4218–4221, July 1981
Biochemistry

A nucleotide change at a splice junction in the human β-globin gene is associated with β^0-thalassemia

(human β-globin/splicing/β^0-thalassemia)

MICHAEL BAIRD*, CATHERINE DRISCOLL*, HELEN SCHREINER*, GIUSEPPINA V. SCIARRATTA†, GENNARO SANSONE†, GULZAR NIAZI‡, FRANCESCO RAMIREZ§, AND ARTHUR BANK*

*Columbia University, College of Physicians and Surgeons, Departments of Medicine, Pediatrics, and Human Genetics and Development, 701 West 168th Street, New York, New York 10032; †Centro della Microcitemia, Ospedali Galliera, Via Volta 8, Genoa, Italy; ‡International Health Resource Center of Hawaii, International Reference and Research Center for Blood Diseases, Kauai Preventive Medicine Program, Lihue, Hauai, Hawaii; and, §Rutgers University Medical School, Department of Obstetrics and Gynecology, Piscataway, New Jersey

Communicated by Elvin A. Kabat, April 6, 1981

ABSTRACT **β^0-Thalassemia is a heterogeneous group of disorders associated with absence of β-globin. In a survey of DNAs from patients with β^0-thalassemia of diverse ethnic origins, a change at the splice junction at the 5′ end of the large intervening sequence (IVS 2) of the human β-globin gene has been found in one patient of Italian and another two of Iranian ethnic origins. The enzyme *Hph* I recognizes a change at this site and generates a larger-than-normal fragment of DNA, which hybridizes specifically to a β-globin IVS 2 probe. No other changes in β-globin gene DNA structure or organization are detectable by extensive restriction endonuclease analysis. The enzyme *Hin*fI which recognizes a sequence beginning three nucleotides from the 5′ end of the IVS 2 splice junction, produces normal fragments and localizes the defect to a G-G-T sequence at the 5′-end IVS 2 splice junction. This sequence is known to be remarkably conserved in all globin genes from many species and in most other genes examined to date. Thus, in at least some β^0-thalassemia patients, the β^0-thalassemia defect is associated with a nucleotide change at a splice junction. These patients provide unique examples of naturally occurring defects in splice junctions of eukaryotic genes associated with absence of specific gene function.**

The β^0-thalassemias are a heterogeneous group of disorders all characterized by absent β-globin synthesis (1–3). At the mRNA level, the β^0-thalassemias are characterized by one of the following: (*i*) the absence of β-globin mRNA; (*ii*) the presence of β-globin mRNA that is abnormal in structure, usually in reduced amounts; or (*iii*) an unspecified defect in β-globin mRNA in patients from the Ferrara region of Italy (4–10). Gene mapping studies to date, using many different restriction enzymes, have failed to reveal deletions or other alterations in nucleotide sequence in the majority of β^0-thalassemia patients examined (11–14). Deletion of the 3′ end of the β-globin gene has been described in a few patients with β^0-thalassemia of Indian extraction (13, 14). In addition, a nonsense mutation at codon 17 has been reported by analyzing the β-globin cDNA structure of a patient with β^0-thalassemia (15). To specifically evaluate splice junctions in the β-globin gene in patients with β^+- and β^0-thalassemia, we utilized a cloned probe representing the large intervening sequence (IVS 2) of the β-globin gene (β-globin IVS 2 probe). Using this probe, we found three individuals homozygous for β^0-thalassemia who generate a unique fragment of DNA that is 100 base pairs longer than the normal fragment with the enzyme *Hph* I. This enzyme recognizes the 5′-end IVS 2 splice junction in the β-globin gene (16, 17); we further localized the nucleotide change to one of three nucleotides of a G-G-T sequence at the 5′-end IVS 2 splice junction using *Hin*fI; the G-G-T sequence is completely conserved in all globin genes from all species studied to date and in most other genes (18, 19). We conclude that this change in nucleotide sequence at the 5′ end of the β-globin IVS 2 splice junction is most likely the cause of the β^0-thalassemia genotype in these individuals.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

All three patients with homozygous β^0-thalassemia found to have DNA changes had the clinical features of severe homozygous β-thalassemia and parents with classical high A_2 β-thalassemia trait. DNA was prepared from peripheral blood samples as described (20). Restriction endonuclease analysis, transfer of DNA fragments to nitrocellulose filters, hybridization to ^{32}P-labeled DNA probes, and radioautography were performed as described (20). The specific probes used were a 4.4-kilobase (kb) *Pst* I cloned DNA fragment (4.4-kb probe), a 0.92-kb β-globin IVS 2 probe, and a 0.95-kb δ-globin IVS 2 probe. Each of these probes were obtained by appropriate cleavage of pBR322 plasmids containing these fragments (21, 22). Hβ1S was obtained from T. Maniatis; this is the pBR322 plasmid containing the 4.4-kb human β-globin gene cloned fragment. The δ- and β-globin IVS 2-containing pBR322 plasmids were prepared in our laboratory by isolation of *Eco*RI–*Bam*HI digestion fragments from the δ- and β-globin IVS 2 and subsequent cloning into pBR322 (22). Each of the probes was isolated from pBR322 after growth and labeled by nick translation to specific activities between 1×10^8 and 8×10^8 cpm/μg (20).

RESULTS

Specificity of β- and δ-Globin IVS 2 Probes. Specificity was demonstrated by hybridizing the β- and δ-globin IVS 2 probes to *Eco*RI- and *Eco*RI/*Bam*HI-digested DNA from normal individuals. Studies (1, 11, 20) have indicated that the 5.2-kb *Eco*RI fragment contains the β-globin IVS 2 and the 2.25-kb *Eco*RI fragment contains the δ-globin IVS 2 (Fig. 1*A*); in this study (Fig. 1*B*), the β-globin IVS 2 probe specifically hybridized to the 5.2-kb fragment, and the δ-globin IVS 2 probe hybridized to the 2.25-kb fragment, indicating their specificity. Further, when normal DNA was cleaved with *Eco*RI and *Bam*HI, the appropriately sized IVS 2-containing fragment was hybridized with its specific probe. In the case of the δ-globin IVS 2, a 0.95-kb fragment was hybridized, whereas with β-globin IVS 2, a 0.92-kb fragment was hybridized (data not shown). These studies indicate that the β-globin IVS 2 probe only hybridizes with

Abbreviations: IVS, intervening sequence; kb, kilobase.

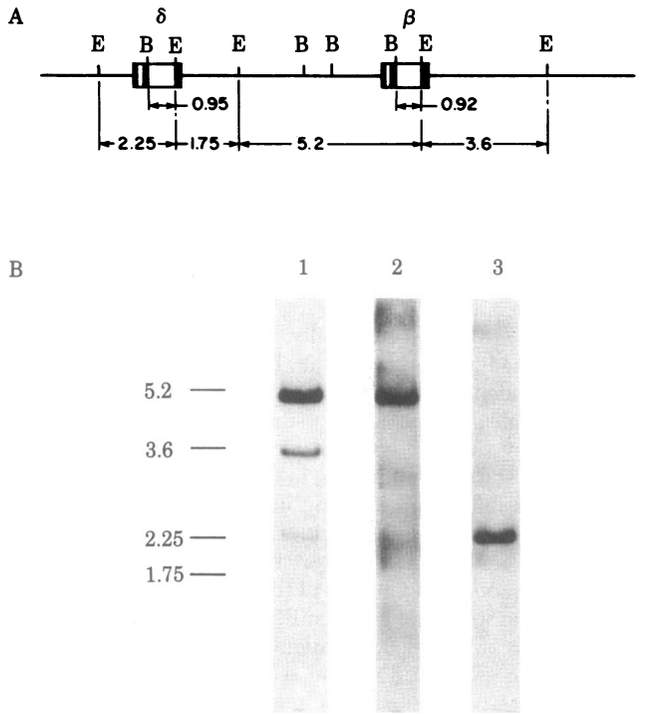

FIG. 1. Specificity of β- and δ-globin IVS 2 probe. (*A*) Restriction map of δ- and β-globin genes showing *Eco*RI (E) and *Bam*HI (B) sites. The rectangles are the structural genes, in which the dark areas are the coding sequences, and the light areas are the IVSs. The sizes of the *Eco*RI fragments and the *Bam*HI–*Eco*RI fragments containing IVS 2 of the δ- and β-globin genes are shown in kb. (*B*) Radioautographs of *Eco*RI-digested normal human DNA, subjected to electrophoresis and hybridization with ^{32}P-labeled DNA probes. The only difference in the three lanes is the origin of the probe (see text). Lanes: 1, 4.4-kb probe; 2, β-globin IVS 2 probe; 3, δ-globin IVS 2 probe. The β-globin IVS 2 hybridizes only to the 5.2-kb fragment, and the δ-globin IVS 2 hybridizes to the 2.25-kb fragment alone, indicating their specificity.

fragments of DNA containing the β-globin IVS 2 and not with other DNA fragments containing either β-globin structural genes alone, δ- or γ-globin gene sequences, or other unrelated DNA sequences.

A Change in *Hph* I Cleavage Site in the β-Globin IVS 2 in Three Patients Homozygous for β^0-Thalassemia. The enzyme *Hph* I is known to cleave at the 5′ end of β-globin IVS 2 (17) (Fig. 2*A*). A single fragment, 0.9-kb in size, was generated from β-globin IVS 2 in normal DNA when this cleavage took place and the β-globin IVS 2 probe was used (Fig. 2). In three patients homozygous for β^0-thalassemia, one of Italian and another two of Iranian ethnic origins, a new fragment, 1.0-kb in size, was found, consistent with the loss of an *Hph* I site at the 5′ end of the β-globin IVS 2 (Fig. 2). Loss of an *Hph* I site 3′ to IVS 2 would be expected to generate a 1.2-kb fragment (Fig. 2*A*). When the three β^0-thalassemia DNAs showing the abnormal 1.0-kb *Hph* I fragment were cleaved with both *Hph* I and *Bam*HI, a 0.95-kb fragment was generated, confirming that the altered *Hph* I sequence is at the 5′ end of the β-globin IVS 2, rather than at the 3′ end (data not shown). If the 3′-end *Hph* I site in β-globin IVS 2 had been affected, this double digestion would have resulted in no significant change in the size of the fragments from that with *Hph* I alone (Fig. 2). Five patients with homozygous β^+-thalassemia and 10 other patients homozygous for β^0-thalassemia produced normal-sized 0.9-kb fragments with *Hph* I (Fig. 2).

Analysis with Other Restriction Enzymes. Cleavage of the β^0-thalassemia DNAs containing the abnormal *Hph* I pattern

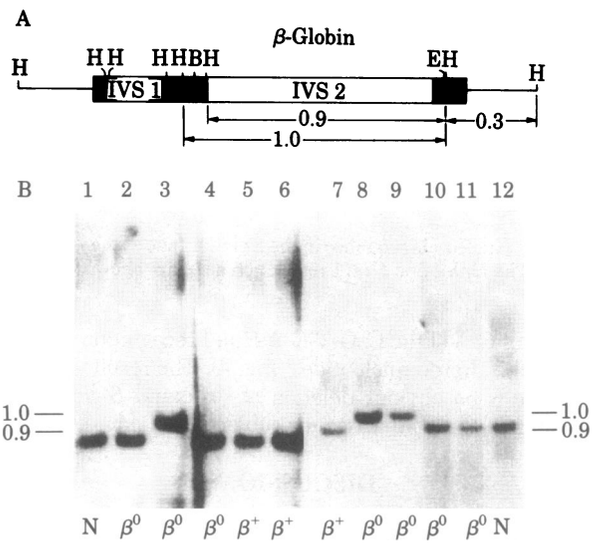

FIG. 2. *Hph* I cleavage of normal and β-thalassemia DNAs. (*A*) Restriction map of the *Hph* I (H), *Bam*HI (B), and *Eco*RI (E) sites in the β-globin gene. The dark areas are the coding regions of the gene. The sizes of *Hph* I fragments are shown in kb. A 0.9-kb fragment is expected with the β-globin IVS 2 probe. (*B*) Radioautographs of *Hph* I-digested DNAs treated as described and hybridized to a ^{32}P-labeled β-globin IVS 2 probe. The source of the DNAs used, normal (N) and β^+- and β^0-thalassemia patients, is indicated below. Lane 3 contains the DNA from the β^0-thalassemia patient from Northern Italy described in the text; lanes 8 and 9 are DNAs from two Iranian patients.

with other enzymes, such as *Eco*RI alone, *Eco*RI and *Bam*HI, *Pst* I, *Asu* I, *Hae* III, and *Hinf* I, showed no changes from normal (Fig. 3). The presence of a normal *Hinf*I pattern with the β-globin IVS 2 probe was particularly informative because this enzyme recognizes the nucleotides that are three to seven bases away from the 5′ end of the IVS 2 (Figs. 3 and 4). Thus, the normal *Hinf*I patterns in the three affected β^0-thalassemia subjects indicate that these nucleotides are intact and eliminate the

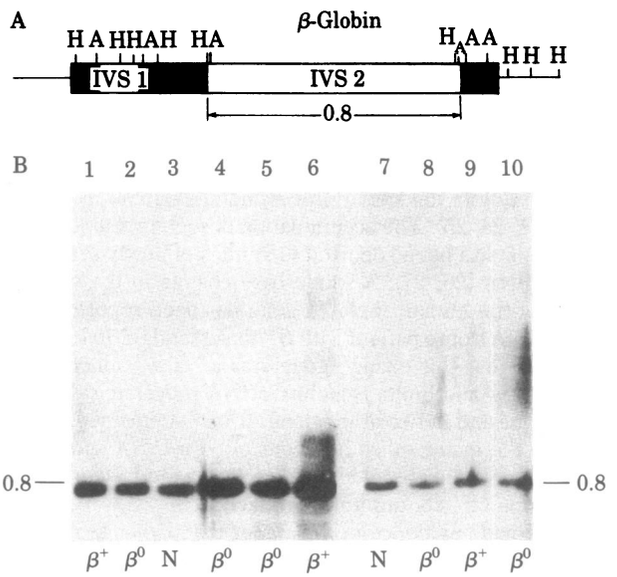

FIG. 3. Restriction analysis of normal (N) and β^+- and β^0-thalassemia DNAs with *Hinf* I (H) and *Asu* I (A). (*A*) Restriction map of the β-globin gene. The dark areas are the coding regions of the gene. The *Hinf* I and *Asu* I fragments expected in normal DNA with the β-globin IVS 2 probe are 0.8 kb. DNA in lanes 1–6 are cleaved with *Hinf* I and in lanes 7–10 with *Asu* I. Lanes 4 and 8 contain DNA from the β^0-thalassemia patient from Italy with an abnormal *Hph* I fragment.

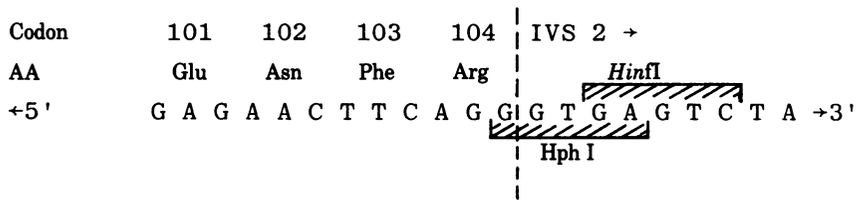

FIG. 4. Nucleotide sequence in the β globin gene at the junction of coding sequence and IVS 2. The vertical line indicates the 5'-end splice junction of IVS 2. The *Hph* I and *Hinf* I restriction sites are shown by brackets and diagonal lines. AA, amino acid.

nucleotides G-A of the G-G-T-G-A *Hph* I recognition sequence as being the altered nucleotides (Fig. 4). The results also show that there is no obvious deletion of structural β-globin gene sequences in the β^0-thalassemia subjects studied.

DISCUSSION

β^0-Thalassemia is a heterogeneous group of disorders characterized by absent β-globin synthesis. Deletion of β-globin structural genes and the presence of a nonsense mutation in the β-globin structural gene have been demonstrated to date as two mechanisms for this disorder (13–15). We now report a change in the nucleotide sequence at the splice junction at the 5' end of IVS 2 of the human β-globin gene in some patients with β^0-thalassemia. We have shown that the usual *Hph* I site at the 5' end of β-globin IVS 2 is missing in these individuals. The nucleotide defect is further localized by two additional observations. First, the fact that *Hinf*I digestion reveals a normal pattern (Fig. 3) indicates that nucleotides three to seven bases from the 5' end of β-globin IVS 2 are intact (Fig. 4). Second, an abnormal hemoglobin, hemoglobin Camperdown (β^{104} Arg→Ser), has been described (23). The replacement of arginine by serine at position 104 in this hemoglobin (Fig. 4) requires a change in the third base of the arginine codon AGG to either a T or a C because the two serine codons are AGT and AGC. The production of hemoglobin Camperdown in significant amounts shows that these changes in sequence just 5' to the splice junction of β-globin IVS 2 are unlikely to cause a β^0-thalassemia defect. A change from AGG to AGA, which also codes for arginine, cannot be ruled out. Therefore, it is possible that the base change in the G-G-T sequence at the 5'-end splice junction is in the G preceding the splice site. However, it is more likely that this change is in either the G or T at the 5' end of IVS 2.

We and others have suggested that β^0- and β^+-thalassemias might be due to defects in processing of β-globin mRNA precursors (2, 24, 25). The accumulation of such precursors in β^+-thalassemia has been reported in studies of newly synthesized nuclear RNA (26, 27). A single base change in the small IVS (IVS 1) of the human β-globin gene has been reported in one cloned DNA from a patient with β^+-thalassemia (28). It has been suggested that this change generates a "false" splicing signal within IVS 1 and limits β-globin mRNA processing. In this β-globin gene and in two other clones from patients with β^+-thalassemia, the major splice junctions of both IVS 1 and IVS 2 of the β-globin gene are intact by direct nucleotide sequence determination (22, 28; unpublished data).

We provide evidence in this paper that a change in the nucleotide sequence at the 5'-end splice junction of β-globin IVS 2 may be the cause of the underlying defect in β-globin synthesis in some individuals with β^0-thalassemia and may completely prevent normal β-globin synthesis. The presence of a single *Hph* I fragment suggests homozygosity for the defect at the splice junction in each of the three affected patients. This is in contrast to evidence for heterozygosity in the gene defects in β^0-thalassemia patients with deletions at the 3' end of the β-globin gene (13, 14). Additionally, we find apparently similar defects in one Italian and two other Iranian patients, most likely caused by independent mutations, although unusual transmission of a single gene remains possible. Pedigree analysis of the Italian family suggests no evidence of consanguinity in three generations studied; inbreeding in the affected populations remains a possibility and is being investigated. Sequencing of these genes should precisely define the nucleotide defect at the 5' end of IVS 2 and clarify other structural changes and, perhaps, the genetic relationships of these genes.

Functional studies of globin mRNA metabolism in the bone marrow cells of the affected patients, which may demonstrate the lack of β-globin mRNA and accumulation of abnormal β-globin mRNA precursors in the nucleus of these cells, will be required to definitively relate the structural abnormality we describe to the functional absence of β-globin. The clones containing these abnormal β-globin genes should provide useful materials for studying the relationship between changes in gene structure at splice junctions and gene function with *in vitro* and *in vivo* systems for gene transcription and RNA processing. The new restriction fragment generated by *Hph* I also should facilitate prenatal diagnosis of fetuses-at-risk for β^0-thalassemia having this defect and an analysis of the population genetics related to this gene.

The G-G-T sequence at the splice junction at the 5' end of IVS 2 is conserved in all globin genes studied in several species and in most other genes (18, 19, 29). The remarkable conservation of this sequence suggests its functional significance and makes it extremely unlikely that the change at the 5' end of the IVS 2 in these patients with β^0-thalassemia represents a polymorphism unrelated to the underlying defect in β-globin synthesis, although this possibility formally exists. Appropriate expression of globin genes or globin cDNAs using simian virus 40 recombinants in monkey cells has been shown to require intact splice junctions for function (30–33). Our results demonstrate the *in vivo* significance of intact splice junctions for normal gene function and are unique examples of naturally occurring nucleotide defects at these junctions. A similar defect involving a G-to-A change at the 5' end of IVS 2 in a cloned β^0-thalassemia gene has been reported recently in abstract form (34).

We would like to thank Ms. Una T. Collins for preparing this manuscript. M.B. and C.D. are postdoctoral fellows supported by grants from the National Institutes of Health (AM-07373 and GM-07419, respectively). This work was made possible by grants from the National Institutes of Health (AM-25274), the National Science Foundation (24120), the Cooley's Anemia Foundation (312-2581), and the March of Dimes–Birth Defects Foundation (NF-1-664).

1. Bank, A., Mears, J. G. & Ramirez, F. (1980) *Science* **207**, 486–493.
2. Bank, A. (1978) *Blood* **51**, 369–384.
3. Bunn, H. F., Forget, B. G. & Ranney, H. M. (1977) *Hemoglobinopathies* (Saunders, Philadelphia), p. 28.

4. Forget, B. G., Benz, E. J., Skoulchi, A., Baglioni, C. & Housman, D. (1974) *Nature (London)* **247**, 379–385.
5. Ramirez, F., Starkman, D., Bank, A., Kerem, H., Cividalli, G. & Rachmilewitz, E. A. (1977) *Blood* **266**, 231–238.
6. Kan, Y. W., Holland, J. P., Dozy, A. M. & Varmus, H. E. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 5140–5144.
7. Ramirez, F., O'Donnell, J. V., Marks, P. A., Bank, A., Musumeci, S., Schiliro, G., Pizzarelli, G., Russo, G., Luppis, B. & Gambino, R. (1976) *Nature (London)* **263**, 471–480.
8. Ottolenghi, S., Comi, B., Giglioni, B., Williamson, R., Vullo, C., Del Senno, L. & Conconi, F. (1977) *Nature (London)* **266**, 231–240.
9. Old, J. M., Proudfoot, N. J., Wood, W. G., Longley, J. I., Clegg, J. B. & Weatherall, D. J. (1978) *Cell* **14**, 289–298.
10. Benz, E. J., Forget, B. G., Helliman, D. G., Cohen-Solal, M., Pritchard, J., Cavallesco, C., Prensky, W. & Housman, D. (1978) *Cell* **14**, 299–312.
11. Mears, J. G., Ramirez, F., Leibowitz, D. & Bank, A. (1978) *Cell* **15**, 15–23.
12. Bank, A., Mears, J. G., Ramirez, F., Burns, A. L. & Feldenzer, J. (1979) in *Eukaryotic Gene Regulation*, eds. Axel, R., Maniatis, T. & Fox, F. C. (Academic, New York), pp. 355–366.
13. Flavell, R. A., Bernards, R., Kooter, J. M. & deBoer, R. (1979) *Nucleic Acids Res.* **6**, 2749–2760.
14. Orkin, S. H., Old, J. M., Weatherall, D. J. & Nathan, D. G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2400–2404.
15. Chang, J. C. & Kan, Y. W. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2886–2889.
16. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
17. Lawn, R. M., Efstradiatis, A., O'Connell, C. & Maniatis, T. (1980) *Cell* **21**, 647–651.
18. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
19. Efstradiatis, A., Tosakany, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. J., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, D., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.
20. Mears, J. G., Ramirez, F., Leibowitz, D., Nakamura, F., Bloom, A., Konotey-Ahulu, F. & Bank, A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1222–1226.
21. Rodriquez, R. L., Bolivar, F., Goodman, H. M., Boyer, H. W. & Betlach, M. (1976) in *Molecular Mechanisms in the Control of Gene Expression*, eds. Nierlich, D. P., Rutter, W. J. & Fox, C. F. (Academic, New York), pp. 471–478.
22. Burns, A. L., Spence, S., Kosche, K., Ramirez, F., Mears, J. G., Schreiner, H., Miller, C., Baird, M., Leibowitz, D., Giardina, P., Markenson, A. & Bank, A. (1981) *Blood* **57**, 140–146.
23. Wilkinson, T., Chua, C. G., Correll, R. W., Robin, H., Exner, T., Lee, K. M. & Kronenberg, H. (1975) *Biochim. Biophys. Acta* **393**, 195–200.
24. Kacian, D. L., Gambino, R., Dow, L. W., Grossbard, E., Natta, C., Ramirez, F., Spiegelman, S., Marks, P. A. & Bank, A. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 1886–1890.
25. Nienhuis, A. W., Turner, P. & Benz, E. J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3960–3964.
26. Kantor, J. A., Turner, P. H. & Nienhuis, A. W. (1980) *Cell* **21**, 149–158.
27. Maquat, L. E., Kinniburgh, A. J., Beach, L. R., Honig, G. R., Lazerson, J., Ershler, W. B. & Ross, J. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4287–4291.
28. 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. & Weissman, S. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2455–2459.
29. Gannon, F., O'Hare, K., Perrin, F., LePennce, J. P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Carri, B. & Chambon, P. (1979) *Nature (London)* **278**, 428–434.
30. Leder, P., Hansen, J. N., Konkel, D., Leder, A., Nishioka, Y. & Talkington, C. (1980) *Science* **209**, 1336–1342.
31. Hamer, D. H. & Leder, P. (1979) *Cell* **18**, 1299–1309.
32. Hamer, D. H. & Leder, P. (1979) *Nature (London)* **281**, 35–42.
33. Mulligan, R., Horward, E. & Berg, P. (1979) *Nature (London)* **277**, 108–112.
34. Shander, M., Van der Woude, S., Proudfoot, N. & Maniatis, T. (1981) *J. Supramol. Struct. Cell. Biochem. Suppl.* **5**, 229.