Proc. Natl. Acad. Sci. USA
Vol. 76, No. 6, pp. 2886–2889, June 1979
Genetics

β^0 thalassemia, a nonsense mutation in man

(cDNA/reverse transcriptase/nucleic acid sequencing/amber mutation)

JUDY C. CHANG AND YUET WAI KAN

Howard Hughes Medical Institute Laboratory and Hematology Service, San Francisco General Hospital, and Department of Medicine and Laboratory Medicine, University of California, San Francisco, California 94143

Communicated by Arno G. Motulsky, April 5, 1979

ABSTRACT **We determined the complete nucleotide sequence of the 5′ noncoding region and the first 74 amino acids of the nonfunctional β-globin mRNA in a patient with homozygous β^0 thalassemia. We identified the molecular defect as a single nucleotide substitution in the coding region of the mRNA. At the position corresponding to amino acid 17, replacement of an adenine by a uracil changes the triplet AAG, which codes for lysine in the normal β chain, to an amber termination codon, UAG. This type of β^0 thalassemia represents an example of a nonsense mutation in man.**

β^0 thalassemia is characterized by the absence of β-globin chain synthesis in the homozygous state. In the majority of cases the β-globin structural gene is present; the amount of β-globin mRNA varies among different patients (1–10). In a Chinese patient with homozygous β^0 thalassemia, we have previously shown by cDNA·RNA hybridization and RNA fingerprint analysis that the authentic β-globin mRNA was present but was not translated *in vivo* or *in vitro* (11). This type of mRNA is of particular interest because the defect in globin synthesis most likely results from structural abnormality of this mRNA. The absence of β-globin synthesis could arise from abnormalities in the 5′ noncoding region of the mRNA, affecting initiation of protein synthesis. It could also be due to a nonsense mutation in the early coding region, resulting in premature termination of the globin chain. To detect the lesion, we analyzed the sequence of this patient's β-globin mRNA.

By labeling the 3′ end of the full-length single-stranded cDNA as described (12), we were able to derive the sequence of the β-globin mRNA from its 5′ end to the first 10 amino acids. However, we were unable to obtain the β-globin coding sequence unambiguously by analyzing *Hae* III-digested fragments of oligo(dT)-primed cDNAs because of contamination by α and γ sequences which constituted more than 90% of the mRNAs in this patient. Therefore, our strategy for sequencing the coding region involved the preparation of polydeoxynucleotides with sequences complementary to regions of β-globin mRNA. These polydeoxynucleotides were used as primers in sequencing the mRNA with the modified chain terminator method of Sanger *et al.* (13). Alternatively, we synthesized the β-specific cDNA by extending the β-specific primer to the 5′ end of the β-globin mRNA from the patient's total mRNAs with reverse transcriptase. We then digested the cDNA with *Hae* III and determined the sequence of these fragments by the method of Maxam and Gilbert (14).

MATERIALS AND METHODS

Source of Globin mRNA. RNA was extracted from the reticulocyte of a Chinese patient with homozygous β^0 thalassemia with nonfunctional β-globin mRNA, and the poly(A)-rich RNA was purified on an oligo(dT)-cellulose column as described (11). As nonthalassemic control, poly(A)-rich RNA was also prepared from a patient with sickle cell anemia.

Determination of Sequence of the 5′ Noncoding Region of the mRNA. We used the method described for determining the nucleotide sequence of the 5′ noncoding regions of human globin mRNAs (12, 15). Briefly, we used oligo(dT) to prime full-length single-stranded cDNA synthesis from the mRNAs with reverse transcriptase, and labeled the 3′ ends of the cDNAs with [α-^{32}P]GTP (New England Nuclear, 300 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels) by using terminal deoxynucleotidyltransferase. We then digested the cDNAs with *Hae* III, isolated the labeled 3′-end β-globin cDNA fragment on a 20% polyacrylamide gel in 7 M urea, and determined its sequence, which corresponded to the 5′ sequence of the mRNA, with the method of Maxam and Gilbert (14).

Preparation of DNA Primers Specific for β-Globin mRNA. The recombinant plasmid JW 102, which contains human β-globin cDNA (16) was prepared and the plasmid DNA was purified by the procedure of Bolivar *et al.* (17). Containment was according to National Institutes of Health guidelines.

We then isolated the DNA fragments containing the globin sequences by digesting the plasmid DNA with restriction endonucleases (New England BioLabs) under the conditions specified by the manufacturers. Plasmid DNA (1 mg) was first digested with 2500 units of *Hha* I and the digestion products were extracted with phenol and precipitated in ethanol. The single fragment that contained the entire β-globin cDNA insert was isolated by discontinuous electrophoresis on a 0.8% preparative agarose gel according to the method of Polsky *et al.* (18); 200 μg of this fragment was digested with 1000 units of *Bam*HI and 500 units of *Eco*RI. The resultant fragments were again isolated by discontinuous electrophoresis on a 2% agarose gel; 20 μg of the fragment that contained the 5′ portion of globin sequence was further digested with 40 units of *Hae* III and 100 units of *Hin*fI. The fragments were separated by electrophoresis on a 5% polyacrylamide slab gel in 50 mM Tris borate, pH 8.3/1 mM EDTA. DNAs were visualized by staining with 0.02% methylene blue in 0.4 M NaOAc (pH 4.5) for 1 hr at room temperature, and the fragments were eluted as described (12).

Determination of mRNA Sequence with Chain-Terminator Method. The sequence of the mRNA from the primer binding site toward the 5′ end was determined by the method of Sanger *et al.* (13) with some modifications. Hybridization mixture (10 μl) containing 0.3 pmol of the mRNA and 1.5 pmol of the primer, 20 mM Tris·HCl at pH 7.4, and 100 mM NaCl was heated at 100°C for 3 min to denature the double-stranded primer and then incubated at 65°C for 30 min to allow the mRNA to anneal with the complementary strand of the DNA primer. Synthesis was performed in four separate reverse transcriptase reactions, each with one of the four chain terminators, 2′,3′-dideoxy ATP, CTP, GTP, and TTP, to generate partially elongated DNA products. Five microliters of the reaction mixture contained 2 μl of the hybridization mixture, 50

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.

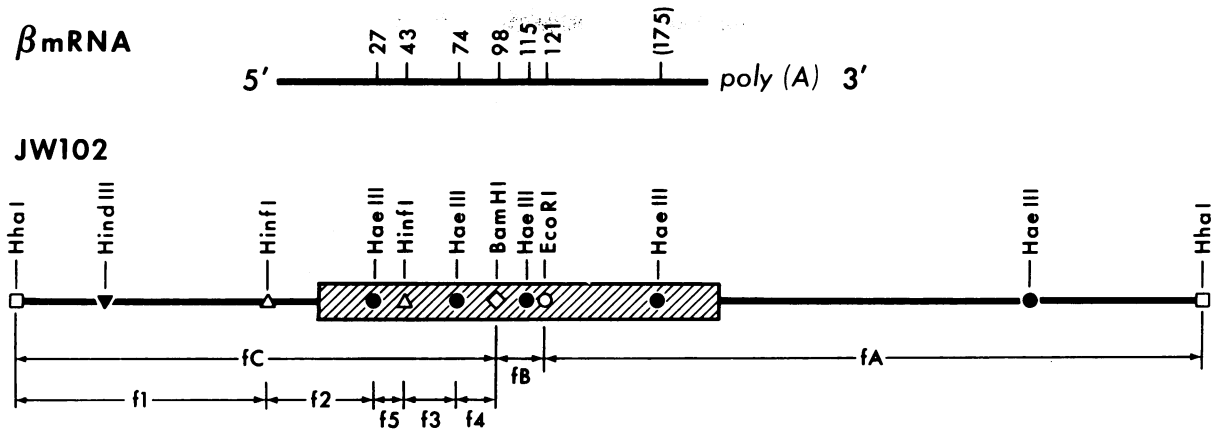

FIG. 1. Restriction map of the 1.7-kilobase *Hha* I fragment of the plasmid JW 102 containing the human β-globin cDNA insert (shaded area). Numbers on the β mRNA indicate positions of amino acid in the β globin. Restriction endonuclease sites on the β globin cDNA were derived from nucleotide sequence data of Marotta *et al.* (20).

mM Tris·HCl (pH 8.3), 10 mM $MgCl_2$, 10 mM dithiothreitol, dATP, dGTP, and TTP each at 25 μM, [α-^{32}P]dCTP (Amersham, 300 Ci/mmol) at 1 μM, 300 units of reverse transcriptase per ml, and an appropriate amount of one of the four dideoxynucleoside triphosphates. The mixtures were incubated at 42°C for 15 min. To each reaction mixture, 1 μl of 0.5 mM dCTP was added and incubation was continued for another 15 min. A mixture of dyes in formamide was added, and the partially elongated DNA products were analyzed on a 8% polyacrylamide slab gel as described (13, 19). Electrophoresis was run at 1.2–1.5 kV for 3–4 hr.

Sequence Determination by the Maxam and Gilbert Method. cDNA was first synthesized by extending the β-specific primer to the 5′ end of the β^0 mRNA as follows. A mixture of 50 pmol of the primer and 25 pmol of the mRNA in 50 μl was preannealed as in the chain-terminator sequencing method. The primer was extended to the 5′ end of the mRNA with reverse transcriptase. The 200-μl reaction mixture contained the above 50 μl, 50 mM Tris·HCl (pH 8.3), 10 mM $MgCl_2$, 10 mM dithiothreitol, 100 μg of actinomycin D per ml, dATP, dGTP, dCTP, and TTP at 400 μM each, [α-^{32}P]dCTP at 0.1 Ci/mmol (diluted from 300 Ci/mmol, New England Nuclear), and 400 units of reverse transcriptase per ml. The reaction mixture was incubated at 37°C for 1 hr, and the cDNA was purified as described (12).

The single-stranded cDNA synthesized from the mRNA was digested with *Hae* III and the restriction fragments were separated by electrophoresis on a 5% polyacrylamide gel. Fragments were visualized by autoradiography and isolated by elution from the gel. Each fragment was then labeled at its 5′ end with ^{32}P by using polynucleotide kinase (New England BioLabs) and [γ-^{32}P]ATP (New England Nuclear, 5000 Ci/mmol) as described (12), and its sequence was determined by the technique of Maxam and Gilbert (14).

RESULTS

Sequence of the 5′ Noncoding Region. With the method of labeling the 3′ end of the cDNA, we derived the entire 5′ noncoding sequence and the sequence corresponding to the first 10 amino acids of the β-globin mRNA in this patient. The sequence, including the initiation codon AUG, was normal (gel not shown). We also determined that the codon for the ninth amino acid was UCU which codes for serine. This finding confirms that the sequence we determined was derived from β- and not from δ-globin mRNA because the latter would have the sequence ACX to code for threonine at this position of the chain.

Isolation of β-Globin DNA Fragments for Use as Primers. When the chimeric plasmid JW 102 was digested with *Hha* I, the entire β-globin cDNA insert was contained in the largest fragment, 1.7 kilobases (16) (Fig. 1). Digestion of this fragment with *Eco*RI and *Bam*HI produced three fragments, fA, fB, and fC. Fragment fB contained the sequence corresponding to

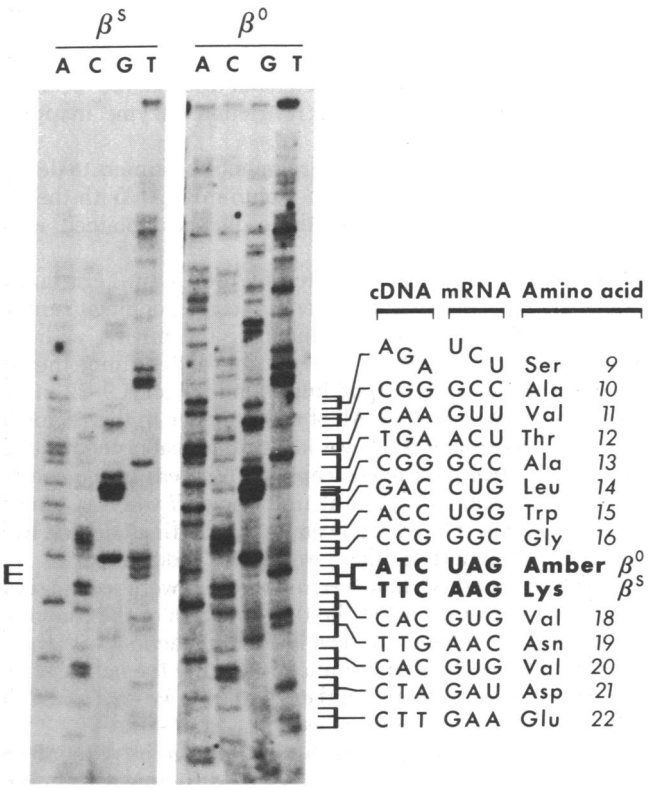

FIG. 2. Sequence of the 5′ end of mRNA by chain-terminator method. β-Globin DNA fragment f5 (amino acids 27–43) was used as primer, and the mRNA from sickle cell anemia (β^S) or β^0 thalassemia was template in the limited synthesis with reverse transcriptase. A, C, G, and T represent four separate reactions containing one of the dideoxynucleoside triphosphates: 40 μM ddATP, 0.4 μM ddCTP, 8 μM ddGTP, or 20 μM ddTTP. Artifacts were observed in this gel, as has been described with this method (13). In addition to the artifact mentioned in the text, at the first nucleotide position of amino acid number 12, a band was seen in all four reactions in both mRNAs; at the three nucleotide positions for amino acid number 13, a "pile-up" of bands occurred. These ambiguities were resolved by using the Maxam and Gilbert method.

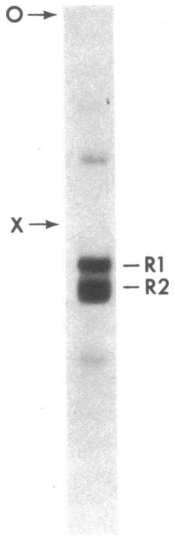

FIG. 3. Separation of *Hae* III-digested cDNA fragments on a 5% polyacrylamide gel. The cDNA was synthesized from β^0 thalassemia mRNA using the β-globin fragment fB (amino acids 98–121) as primer. Arrows indicate positions of origin (O) and xylene cyanol blue marker (X).

amino acids 98–121. Further digestion of the fragment fC with *Hin*fI and *Hae* III produced five fragments. Of these, fragments f5, f3, and f4 contained sequences that corresponded to amino acids 27–43, 43–74, and 74–98, respectively. The partial restriction map surrounding the globin insert in the recombinant plasmid was derived by restriction enzyme mapping technique (gels not shown).

We used these globin DNA fragments as primers to derive the sequence of the patient's β-globin mRNA. With the exception described below, all the sequences obtained were normal.

Abnormal Sequence Detected by the Chain-Terminator Method. When we used fragment f5 (corresponding to amino acids 27–43) as primer in sequencing the mRNA with the chain-terminator method, we were able to determine the nucleotide sequence corresponding to amino acids 9–22 in the β-globin mRNA (Fig. 2). All of the sequence in this region of the β^0 thalassemia mRNA was identical to that of the nonthalassemia control with one exception. Whereas in the control mRNA the codon for lysine at amino acid 17 was AAG, in the β^0 thalassemia mRNA replacement of the first adenine by a uracil resulted in the amber termination codon UAG.

With this method, however, ambiguities were seen in some positions. For example, at the third position of the codon for amino acid 16, a band was seen in all four reactions with the β^S mRNA. Hence, to establish unambiguously the nature of the abnormality we detected in the β^0 mRNA, we confirmed the finding with a different method.

Confirmation of the Amber Mutation by Use of the Maxam and Gilbert Method. We used fragment fB (amino acids 98–121) as primer to extend cDNA synthesis to the 5′ end of the β^0 mRNA. The single-stranded cDNA was then digested with *Hae* III and the fragments were separated on a 5% polyacrylamide gel, (Fig. 3). Two fragments, R1 and R2, were eluted from the gel, labeled at the 5′ ends with ^{32}P by using polynucleotide kinase (12), and further purified on an 8% polyacrylamide gel. Fragment R1 contained the sequence corresponding to amino acids 27–74 of the β-globin and was identical to the normal (gel not shown). Fragment R2 contained a sequence that corresponded to the region from amino acid 27 to the 5′ end of the mRNA (Fig. 4). It confirms the presence

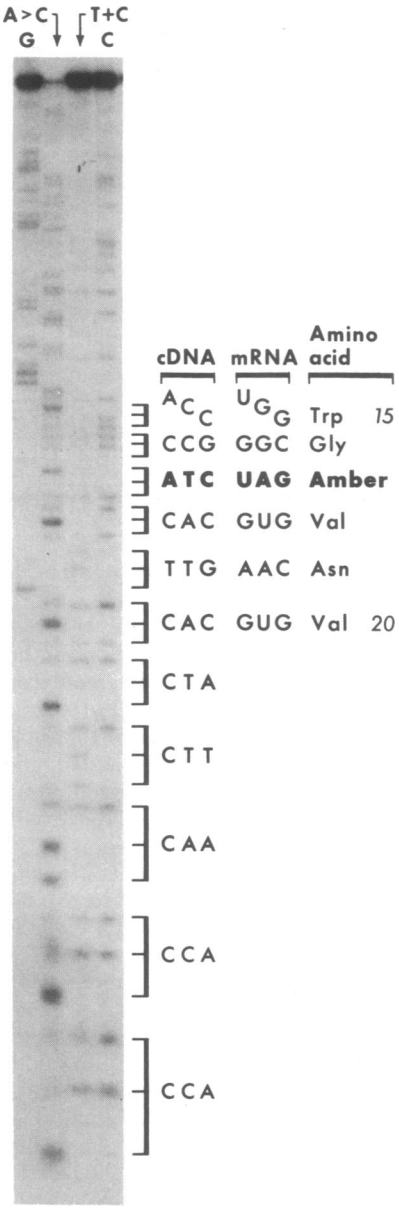

FIG. 4. Sequence of the 5′ end ^{32}P-labeled *Hae* III fragment R2 by chemical degradation method. The partial cleavage products were separated on a 12% polyacrylamide gel in 7 M urea.

of UAG instead of AAG at the position corresponding to amino acid 17.

DISCUSSION

These studies localize the molecular lesion that results in the absence of β-globin chain synthesis in a patient with homozygous β^0 thalassemia. The defect has been identified as a single nucleotide substitution in the coding region of the β-globin mRNA. At the position corresponding to amino acid 17, replacement of the first adenine in the lysine codon AAG by a uracil results in the amber termination codon UAG. This causes premature termination of the β-globin chain in this position. This type of β^0 thalassemia is therefore an example of a nonsense mutation in man.

Studies from several laboratories have indicated that the thalassemia syndromes are a heterogeneous group of disorders in which defective synthesis of the normal globin chain could arise from different molecular mechanisms. These include deletion of globin genes (5, 21–25), unequal crossover events

(26, 27), and termination mutations (28–32). Even in β^0 thalassemia, the underlying molecular defect could be quite diverse. For example, the type of β^0 thalassemia without β-globin mRNA could be due to deletion of the β-globin structural gene or defective DNA transcription or mRNA processing. In the type with detectable β-globin mRNA, in addition to the nonsense mutation described here possible mechanisms could include frameshift mutations and mutations affecting the initiation codon or other ribosomal binding sites. The Ferrara type of β^0 thalassemia (33) with inducible β-globin synthesis may arise from a different mechanism. It is likely that other nonsense mutations will be found in thalassemia. These mutations may occur at different positions along the globin chains. For example, in the sequence shown in Fig. 2, single nucleotide substitutions at amino acid 15 (UGG to UAG) and 22 (GAA to UAA) would produce termination codons. Hence, we propose that, in accordance with the nomenclature used for abnormal hemoglobins, nonsense mutations should be defined by the amino acid position and the codon changed. Thus, the mutant we have described would be designated as $\beta^{17\ AAG\rightarrow UAG}$ or $\beta^{17\ UAG}$.

The small amount of β-globin mRNA present in the patient we studied can be due to one or both of the following. First, the patient could be doubly heterozygous for two types of β^0 thalassemia, one with early termination and the other with no production of β-globin mRNA. Second, because the globin chain synthesis prematurely terminates at amino acid 17, a large part of the mRNA would not be covered by ribosomes and so would be exposed to nuclease degradation. Hence, measurement of newly synthesized β-globin mRNA in this patient's erythroid precursor cells may provide useful information on the effect of protein synthesis on mRNA turnover in mammalian cells.

At present, no method is available for correcting the primary defect in the DNA in a human genetic disorder. The nature of the defect in this type of β^0 thalassemia suggests that an alternative approach could be explored. It is known that nonsense mutation in eukaryotic cells or viruses can be suppressed *in vitro* by species of suppressor tRNA from yeasts (34–37). A possible way of overcoming the defect in β^0 thalassemia would be to suppress the nonsense mutation with these tRNAs. Alternatively, other conditions which result in suppression of nonsense mutation can also be tested with this patient's erythroid cells.

We thank Dr. Bernard Forget for the plasmid JW 102 and the Office of Program Resources and Logistics (Viral Cancer Program, Viral Oncology, Division of Cancer Cause and Prevention, National Cancer Institute, Bethesda, MD) for the reverse transcriptase. This research was supported by grants from the National Institutes of Health (AM 16666, HL 20985) and the National Foundation–March of Dimes. Y.W.K. is an Investigator of the Howard Hughes Medical Institute.

1. Forget, B. G., Benz, E. J., Jr., Skoultchi, A., Baglioni, C. & Housman, D. (1974) *Nature (London)* **247,** 379–381.
2. Ottolenghi, S., Lanyon, W. G., Williamson, R., Weatherall, D. J., Clegg, J. B. & Pitcher, C. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 2294–2299.
3. Kan, Y. W., Holland, J. P., Dozy, A. M. & Varmus, H. E. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 5140–5144.
4. Tolstoshev, P., Mitchell, J., Langon, G., Williamson, R., Ottolenghi, S., Comi, P., Giglioni, B., Masera, G., Modell, B., Weatherall, D. J. & Clegg, B. J. (1976) *Nature (London)* **259,** 95–98.
5. Ramirez, F., O'Donnell, J. V., Marks, P. A., Bank, A., Musumeci, S., Schiliro, G., Pizzarelli, G., Russo, G., Luppis, G. & Gambino, R. (1976) *Nature (London)* **263,** 471–475.
6. Forget, B. G., Hillman, D. G., Cohen-Solal, M. & Prensky, W. (1976) *Blood* **48,** 998A.
7. Godet, J., Verdier, G., Nigon, V., Belhani, M., Richard, R., Colonna, P., Mitchell, J., Williamson, R. & Tolstoshev, P. (1977) *Blood* **50,** 463–470.
8. Ottolenghi, S., Comi, P., Giglioni, B., Williamson, R., Vullo, G. & Conconni, F. (1977) *Nature (London)* **266,** 231–234.
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., Jr., Forget, B. G., Hillman, D. G., Cohen-Solal, M., Pritchard, J. & Cavallesco, C. (1978) *Cell* **14,** 299–312.
11. Temple, G. F., Chang, J. C. & Kan, Y. W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3047–3051.
12. Chang, J. C., Temple, G. F., Poon, R., Neumann, K. H. & Kan, Y. W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5145–5149.
13. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
14. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
15. Chang, J. C., Poon, R., Neumann, K. H. & Kan, Y. W. (1978) *Nucleic Acids Res.* **5,** 3515–3522.
16. Wilson, J. T., Wilson, L. B., deRiel, J. K., Villa-Komaroff, L., Efstratiadis, A., Forget, B. G. & Weissman, S. M. (1978) *Nucleic Acids Res.* **5,** 563–581.
17. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. L., Heyneker, H. L. & Boyer, H. W. (1977) *Gene* **2,** 95–113.
18. Polsky, F., Edgell, M. H., Seidman, J. G. & Leder, P. (1978) *Anal. Biochem.* **87,** 397–410.
19. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87,** 107–109.
20. Marotta, C. A., Wilson, J. T., Forget, B. G. & Weissman, S. M. (1977) *J. Biol. Chem.* **252,** 5040–5053.
21. Ottolenghi, S., Lanyon, W. G., Paul, J., Williamson, R., Weatherall, D. J., Clegg, J. B., Pritchard, J., Pootrakul, S. & Wong, H. B. (1974) *Nature (London)* **251,** 389–391.
22. Taylor, J. M., Dozy, A. M., Kan, Y. W., Varmus, H. E., Lie-Injo, L. E., Ganesan, J. & Todd, D. (1974) *Nature (London)* **251,** 392–393.
23. Ramirez, F., Natta, C., O'Donnell, J. V., Canale, V., Bailey, G., Sanguensermsri, T., Maniatis, G., Marks, P. & Bank, A. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 1550–1554.
24. Kan, Y. W., Dozy, A. M., Varmus, H. E., Taylor, J. M., Holland, J. P., Lie-Injo, L. E., Ganesan, J. & Todd, D. (1975) *Nature (London)* **255,** 255–256.
25. Ottolenghi, S., Comi, P., Giglioni, B., Tolstoshev, P., Lanyon, W. G., Mitchell, G. J., Williamson, R., Russo, G., Musumeci, S., Schiliro, G., Tsistrakis, G. A., Charache, S., Wood, W. G., Clegg, J. B. & Weatherall, D. J. (1976) *Cell* **9,** 71–80.
26. Baglioni, C. (1962) *Proc. Natl. Acad. Sci. USA* **48,** 1880–1886.
27. Weatherall, D. J. & Clegg, J. B. (1972) *The Thalassaemia Syndromes* (Blackwell, Oxford), 2nd Ed., pp. 172–188.
28. Clegg, J. B., Weatherall, D. J. & Milner, P. F. (1971) *Nature (London)* **234,** 337–340.
29. Milner, P. F., Clegg, J. B. & Weatherall, D. J. (1971) *Lancet* **i,** 729–732.
30. Clegg, J. B., Weatherall, D. J., Contopolou-Griva, I., Caroutsos, K., Poungouras, P. & Tsevresis, H. (1974) *Nature (London)* **251,** 245–247.
31. DeJong, W. W. W., Khan, P. M. & Bernini, L. F. (1975) *Am. J. Hum. Genet.* **27,** 81–90.
32. Bradley, T. B., Wohl, R. C. & Smith, G. J. (1975) *Clin. Res.* **23,** 131A.
33. Conconi, F., Del Senno, L., Ferrarese, P., Menini, C., Borgatti, L., Vullo, C. & Labie, D. (1975) *Nature (London)* **254,** 256–259.
34. Gesteland, R. F., Wills, N., Lewis, J. B. & Grodzicker, T. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4567–4571.
35. Cepecchi, M. R., Vonder Haar, R. A., Capecchi, N. E. & Sveda, M. M. (1977) *Cell* **12,** 371–381.
36. Philipson, L., Anderson, P., Olshevsky, U., Weinberg, R., Baltimore, D. & Gesteland, R. (1978) *Cell* **13,** 189–199.
37. Cremer, K. J., Bodemer, M., Summers, W. P., Summers, W. C. & Gesteland, R. F. (1974) *Proc. Natl. Acad. Sci. USA* **76,** 430–434.