*Proc. Natl. Acad. Sci. USA*
Vol. 80, pp. 2318–2322, April 1983
Genetics

# "Silent" nucleotide substitution in a $\beta^+$-thalassemia globin gene activates splice site in coding sequence RNA

(RNA processing/mRNA/expression vector/gene cloning/mutation)

MERRILL E. GOLDSMITH*, R. KEITH HUMPHRIES, TIMOTHY LEY, AMANDA CLINE, JUDITH A. KANTOR†, AND ARTHUR W. NIENHUIS

Clinical Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20205

*Communicated by Stuart Kornfeld, January 24, 1983*

ABSTRACT **A $\beta^+$-thalassemia globin gene was isolated from the genome of a Black individual by molecular cloning. DNA sequence analysis revealed only a single difference between this gene and the normal human $\beta$-globin gene—adenine is substituted for thymine in the third position of codon 24. Codon 24 in both the normal gene (GGT) and the $\beta^+$-thalassemia gene (GGA) encodes glycine. The function of this $\beta^+$-thalassemia gene was compared to the function of the normal human $\beta$-globin gene in monkey kidney cells by using plasmid expression vectors. The codon 24 substitution activates a 5′ splice site that involves the guanine-thymine dinucleotide present in codon 25, 16 nucleotides upstream from the normal exon 1–intron I boundary. The splice, involving the abnormal 5′ site in codon 25, is completed with the normal 3′ splice site at the end of intron I. This splicing abnormality leads to a 75% decrease in the accumulation of normally processed $\beta$-globin mRNA, thereby causing the $\beta^+$-thalassemia phenotype.**

The $\beta$-thalassemias are hereditary anemias characterized by decreased synthesis of structurally normal $\beta$ globin (1, 2). A quantitative deficiency of $\beta$ globin mRNA has been demonstrated in the bone marrow cells of individuals with severe $\beta$-thalassemia. Molecular studies have shown that $\beta$-thalassemia mutations may reflect gene deletion (1, 3, 4) or, more frequently, point mutations that affect $\beta$-globin RNA synthesis (5, 6), processing (5, 7–12), stability (13), or translation (14–17). $\beta$-Thalassemia mutations are of general interest because specific nucleotide substitutions in such genes may be related to their effect on RNA metabolism in intact bone marrow cells. Analyses of such mutations are likely to yield novel structure–function correlations not readily achieved by other experimental strategies.

The $\beta^+$-thalassemia mutations do not destroy gene function entirely but instead result in decreased globin mRNA production. Studies of bone marrow cells from individuals with such mutations suggest that transcription of the $\beta$-globin genes is usually normal, but abnormal processing of the $\beta$-globin gene transcript or instability of the $\beta$-globin mRNA leads to a quantitative deficiency in $\beta$-globin mRNA (18–21). These functional studies have been supported by cloning and analyzing the sequences of several genes from patients with $\beta^+$-thalassemia (5, 7–9). The abnormally processed RNA generated by one of these defective genes has been studied by using cloned genes in expression systems (10, 11).

We have cloned and studied a $\beta^+$-thalassemia globin gene from a Black individual with severe $\beta$-thalassemia that has been found to be due to a functional defect in mRNA processing (patient 2 of ref. 19). Three globin genes isolated from this individual differed from the normal $\beta$-globin gene by only a single nucleotide; this substitution in codon 24 is "silent" at the level of protein sequence. By using a cellular expression system, we found that the codon 24 substitution leads to activation of a 5′ splice site in coding sequence RNA, thereby causing a quantitative decrease in $\beta$-globin mRNA accumulation. These studies show how a subtle change in the sequence of a gene (a change that superficially appeared to be functionally silent) can have a marked effect on mRNA production.

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

**Molecular Cloning and DNA Sequence Analysis.** A human genomic library was constructed with DNA isolated from bone marrow cells (19). The DNA was partially digested with *Eco*RI, ligated to the *Eco*RI arms of the bacteriophage λ vector Charon 4A, and packaged to form a cloned DNA fragment library (22, 23). Clones containing globin genes were identified by the miniplasmid or πVX screening procedure developed by B. Seed and T. Maniatis (24). After subcloning of appropriate DNA fragments (23), the nucleotide sequences of both strands of the $\beta^+$-thalassemia globin gene were determined by using the partial chemical degradation method of Maxam and Gilbert (25) and the "dideoxy" chain termination method of Sanger *et al.* (26). The single-stranded DNA fragments for the latter procedure were generated by using the bacteriophage M13 mp7 (27). The strategy used for sequence analysis was based upon the published restriction map of the human $\beta$-globin gene (28).

**Globin Gene Expression in Monkey Kidney Cells.** The cellular expression system and the construction of the plasmid expression vector have been recently described (29, 30). Globin RNA, synthesized in COS monkey kidney cells, was characterized by S1 nuclease mapping (31, 32). Uniformly labeled probes were synthesized *in vitro* by using single-stranded DNA obtained from M13 bacteriophage recombinants, each of which contained a fragment from the human $\beta$-globin gene (33). The 5′- or 3′-end-labeled probes were made by incubating isolated DNA fragments with polynucleotide kinase or DNA polymerase I, respectively (34). The conditions for hybridization, S1 nuclease digestion, and electrophoretic separation of the protected DNA fragments have been described (29, 33). Analysis of RNA splicing was also performed by primer extension with reverse transcriptase (12, 24).

## RESULTS

**Molecular Cloning and DNA Sequence Analysis of a $\beta$-Thalassemia Globin Gene.** Three recombinant bacteriophage

Abbreviations: bp, base pair(s); kb, kilobase pair(s); nt, nucleotide(s); SV40, simian virus 40.

* Present address: Clinical Pharmacology Branch, National Cancer Inst., National Inst. of Health, Bethesda, MD 20205.
† Present address: Dept. of Biochemistry, Georgetown Univ., Washington, DC 20007.

containing human β-globin genes were isolated. These represented independent cloning events, because each contained a different combination of *Eco*RI fragments from the β-globin gene region. Two recombinant bacteriophage contained intact β-globin genes and a third contained only the 5′ *Eco*RI fragment. The sequence of this 5′ *Eco*RI fragment was determined from 200 base pairs (bp) 5′ to the transcription start site to the intragenic *Eco*RI site. Only a single nucleotide difference was found on comparison to previously published sequences of normal genes (28, 35) (Fig. 1). Both the normal codon 24 (GGT) and the mutant codon (GGA) code for glycine. The two intact globin genes were also found to contain this substitution. The sequence of one was determined from the intragenic *Eco*RI site to a position 50 bp beyond the normal poly(A) addition site. This region of the $\beta^+$-thalassemia gene was identical to the normal human β globin gene. The genomic DNA of this individual was characterized by restriction enzyme digestion and Southern blot analysis. Homozygosity was demonstrated at four sites (Table 1) that are known to be polymorphic (37–40). These data suggest that our patient may be homozygous for the codon 24 substitution, although exceptions to the correlation of restriction enzyme polymorphisms (haplotype) with specific mutations have already been noted (5).

**Function of the $\beta^+$-Thalassemic Globin in a Cellular Gene Expression System.** The $\beta^+$-thalassemia globin gene was studied by using plasmid expression vectors (29) designed to replicate in the monkey kidney COS cell line (29, 30, 41–43) (Fig. 2*A*). Function of the β-globin gene promoter in such vectors requires the presence of the 72-bp directly repeated DNA sequences found near the SV40 origin of replication (12, 29, 44). All three independently isolated globin genes gave identical results. The concentration of β-globin RNA in COS cells transfected with the thalassemia genes (vβ-thal) was always considerably less than in cells transfected with the normal gene (vβ-nor) (data not shown).

**Quantitative Deficiency of mRNA Transcribed From the β-Thalassemia Globin Gene Is Due to the Substitution in Codon 24.** The fragment containing the $\beta^+$-thalassemia globin gene in vβ-thal contains approximately 5,000 bp of DNA, of which the sequence of only 1,850 bp was determined; 1,250 bp of 5′ flanking DNA and 1,900 bp of 3′ flanking DNA sequences were not characterized. A series of hybrid genes was constructed in the expression vector (Fig. 2*B*) to determine if the quantitative deficiency in β-globin RNA was due to the substitution in codon 24 or to some uncharacterized mutation in flanking DNA.

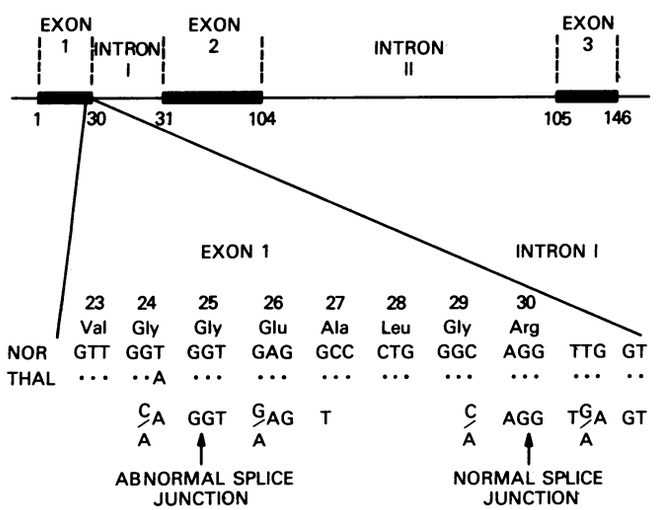


FIG. 1. Nucleotide substitution and position of the activated splice junction in the $\beta^+$-thalassemia globin gene. A general diagram of β-globin gene structure is shown. Three exons encode the 146 amino acids of the polypeptide chain. Below is the amino acid sequence of β-globin and the sequence of the β-globin gene in the region where the single substitution occurs in the $\beta^+$-thalassemia globin gene. Nor, normal sequence; Thal, $\beta^+$-thalassemia sequence. The homology to the 5′ splice consensus sequence (36) and the positions of the splice junctions are indicated below the DNA sequence.

Table 1. Analysis of polymorphic restriction endonuclease sites

| Globin gene | Site | Genotype |
|---|---|---|
| $^{G}\gamma$ | *Hin*dIII | +/+ |
| $^{A}\gamma$ | *Hin*dIII | −/− |
| β | *Ava* II | +/+ |
| 3′ to β | *Bam*HI | +/+ |

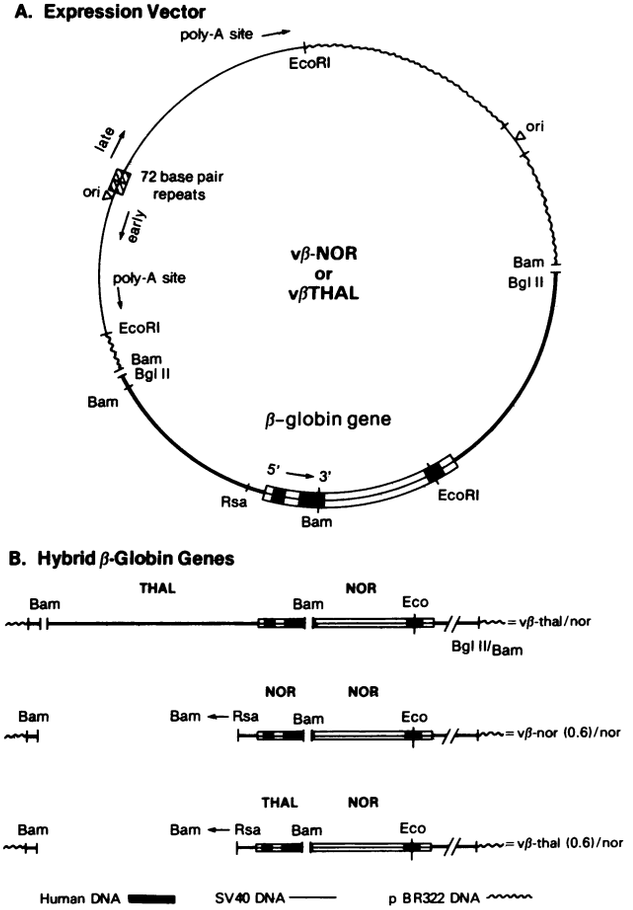


FIG. 2. Diagram of the expression vector and the "hybrid" globin genes used to define the consequences of the nucleotide substitution in codon 24. (*A*) Expression vector. This vector includes two origins of DNA replication (ori) and the promoter enhancer element from the simian virus 40 (SV40) genome (29). The two complete globin genes and the normal human β-globin gene (28) were separately subcloned as 5.0-kilobase (kb) *Bgl* II fragments in the *Bam*HI site of the expression vector. The 5′ end of the third independently isolated partial globin gene was substituted for the corresponding 2.9-kb *Bgl* II/*Eco*RI fragment of the normal human β-globin gene, previously subcloned in the *Pst* I site of the expression vector (29). (*B*) Hybrid globin genes. The vector vβ-thal/nor was obtained by substituting the 5′ end of the $\beta^+$-thalassemia gene (1.9-kb *Bam*HI fragment) for the corresponding fragment in the normal gene. Both the normal and β-thalassemia 1.9-kb *Bam*HI fragments were truncated at the *Rsa* I site 127 bp upstream from the normal start of transcription. The *Rsa* I ends were modified by addition of *Bam*HI linkers and the resulting 0.6-kb *Bam*HI fragments were substituted for the 1.9-kb *Bam*HI fragment in vβ-nor to yield vβ-nor(0.6)/nor and vβ-thal(0.6)/nor, respectively.

These vectors were separately introduced into COS cells. The level of β-globin mRNA present 48 hr after transfection was estimated by S1 nuclease analysis using a uniformly labeled probe that maps the 5′ end of β-globin RNA (Fig. 3*A*). A direct comparison was made between the expression of the β-globin genes in vβ-nor and vβ-thal/nor (Fig. 3*A*, lanes 1 and 2 compared to lanes 3 and 4) and the function of the genes in vβ-nor(0.6)/nor and vβ-thal(0.6)/nor (Fig. 3*A*, lanes 5 and 6 compared to lanes 7 and 8). In each experiment, β-globin mRNA concentration was several fold higher for the normal gene than for the hybrid gene containing the 5′ end of the β-thalassemia globin gene. To estimate the difference in concentration of β-globin RNA molecules transcribed from vβ-nor(0.6)/nor compared to vβ-thal(0.6)/nor more precisely, an analysis was performed with an internal standard. A probe for RNA sequences initiated from the SV40 late promoter in the vector was added to the probe for β-globin RNA sequences (data not shown). Use of the internal standard probe allows correction for variable recovery of RNA during extraction, hybridization, and S1 nuclease analysis. The original radiograph was scanned with a densitometer to measure the relative intensities of the fragments protected by RNA. This analysis revealed the concentration of β-globin RNA in COS cells transfected with vβ-thal(0.6)/nor to be 25% of that in COS cells transfected with vβ-nor(0.6)/nor.

**Substitution in Codon 24 Activates a 5′ Splice Site.** S1 nuclease mapping of the 5′ region of globin RNA transcribed from vectors containing the 5′ end of the β-thalassemia globin gene revealed two protected probe fragments. A 132-nt fragment was identical to the fragment protected by RNA transcribed from vectors containing the normal β-globin gene, and a second fragment was 6 nt shorter (Fig. 3*A*, lanes 3 and 8). The uniformly labeled probe spans the 5′ end of normally initiated β-globin mRNA, but ends 10–11 nt before the splice site at the 5′ end of intron I (33). The 126-nt probe fragment could therefore have arisen because of aberrant initiation of transcription 6 nt downstream from the normal site or because a splice site was utilized within RNA sequences transcribed from exon 1.

The experiment shown in Fig. 3*B* demonstrated that the latter interpretation was correct. A 3′-end-labeled probe derived from the $\beta^+$-thalassemia globin gene was used to precisely map the 5′ splice site by S1 nuclease analysis. Correctly spliced RNA transcribed from vβ-nor(0.6)/nor yielded only a 72-nt band (Fig. 3*B*, lane 1), whereas RNA transcribed from vβ-thal(0.6)/nor protected an additional fragment 57 nt long (Fig. 3*B*, lane 2), indicating utilization of an abnormal 5′ splice site. Aberrantly

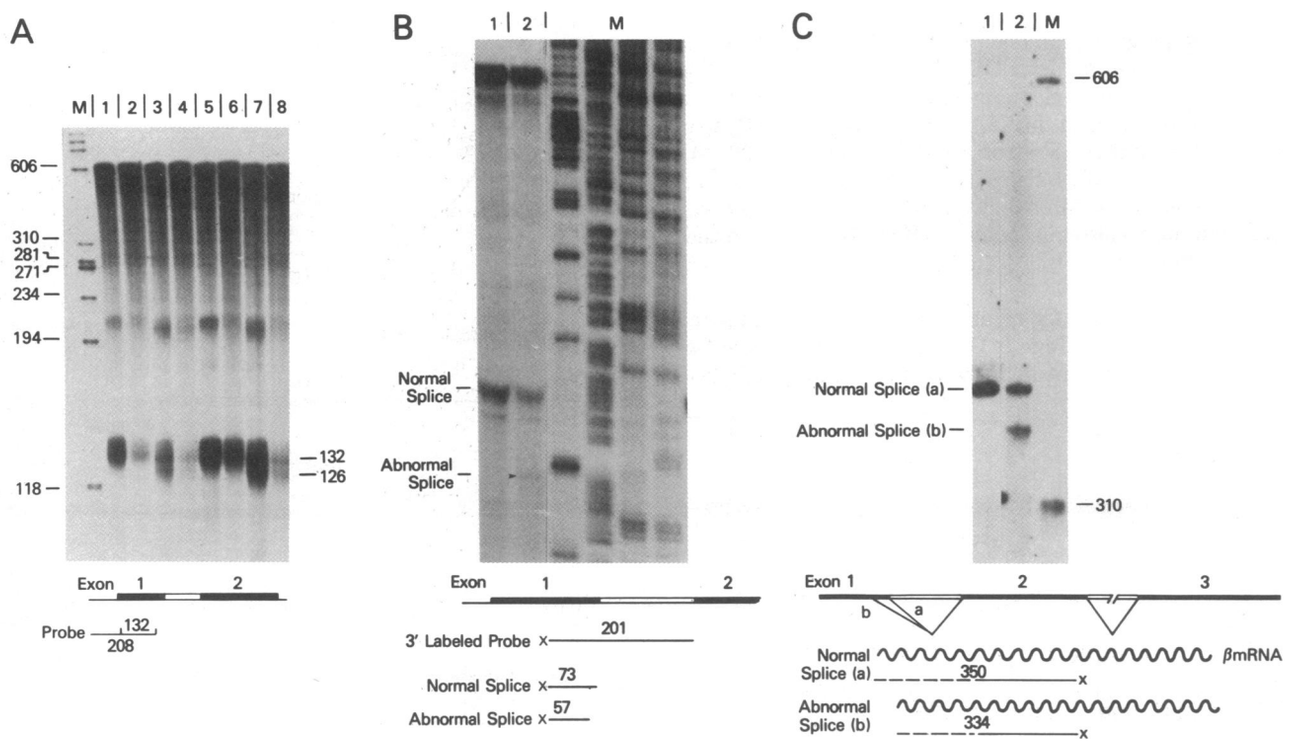


FIG. 3. (*A*) Functional consequences of the codon 24 substitution. Comparison of β-globin RNA concentration in cells transfected with vectors containing each of the hybrid globin genes (Fig. 2). Each RNA sample was annealed with approximately 5 ng of the uniformly $^{32}$P-labeled single-stranded DNA probe (33) shown on the diagram. After exposure to S1 nuclease, the DNA·RNA duplexes were denatured and the protected DNA fragments were resolved on an 8% polyacrylamide sequencing gel. M, marker DNA fragments, with lengths indicated in nucleotides (nt). The odd-numbered lanes show results from analysis of 20 μg of RNA, and the even-numbered lanes show results of analysis of 4 μg of RNA. Vectors studied were as follows: Lanes 1 and 2, vβ-nor; lanes 3 and 4, vβ-thal/nor; lanes 5 and 6, vβ-nor(0.6)/nor; lanes 7 and 8, vβ-thal(0.6)/nor. Note that the $^{32}$P-labeled probe DNA used includes 406 nt of M13 DNA that hybridized with contaminating nonradioactive template DNA and produced the smear of DNA fragments on the upper part of the gel. The additional band at 208 nt represents protection of the entire human DNA portion of the probe by RNA initiated upstream from the normal position (33). (*B*) S1 nuclease mapping using a 3′-end-labeled probe. The 201-bp *Dde* I fragment was isolated from the $\beta^+$-thalassemia gene and $^{32}$P-labeled by incubation with DNA polymerase I, and 20 ng was added to each reaction mixture. Lane 1, 4 μg of RNA from COS cells transfected with vβ-nor(0.6)/nor; lane 2, 20 μg of COS cell RNA from cells transfected with vβ-thal(0.6)/nor. The annealing was at 65°C. The difference in lengths of the two fragments protected by globin RNA transcribed from vβ-thal(0.6)/nor was estimated as 15–16 nt by counting the individual bands in the adjacent sequence ladder. (*C*) Demonstration that the abnormal 5′ splice junction is spliced to a normal 3′ splice junction. A 188-bp *Ava* II/*Bam*HI fragment from exon 2 was $^{32}$P-labeled at the *Bam*HI site with polynucleotide kinase. This primer was annealed with globin RNA and the primer was extended to the 5′ end of the mRNA with reverse transcriptase. The cDNA·RNA hybrids were denatured and the end-labeled cDNA fragments were resolved on polyacrylamide sequencing gels. Lane 1, 20 μg of RNA from COS cells transfected with vβ-nor(0.6)/nor; lane 2, 20 μg of RNA from cells transfected with vβ-thal(0.6)/nor.

spliced RNA molecules of this type were also found in RNA extracted from COS cells transfected with vβ-thal, a vector that contains an intact β-thalassemia globin gene (data not shown).

Primer extension analysis was used to identify the 3′ splice site used with the abnormal 5′ splice site. An *Ava* II/*Bam*HI primer fragment from exon 2 of the β-globin gene was annealed with globin RNA and a cDNA copy of the RNA was synthesized with reverse transcriptase. RNA isolated from COS cells transfected with vβ-nor(0.6)/nor (Fig. 3*C*, lane 1) yielded only the expected 350-nt fragment, whereas RNA from COS cells transfected with vβ-thal(0.6)/nor gave two reverse transcripts (Fig. 3*C*, lane 2). One was 350 nt and the other was approximately 16 nt shorter. This experiment (in conjunction with the previous one) indicated that both the normal and the 5′ splice site in the coding sequence used the same 3′ splice site at the end of intron II. This experiment also provided an accurate estimate of the relative amounts of the two spliced species in steady-state RNA. Sixty-six percent of globin RNA molecules were spliced at the correct site, and 33% were spliced at the codon 25 splice site.

In Fig. 3*C*, very faint (but equivalent) bands derived from unprocessed globin RNA could be seen at 479 nt with RNA transcribed from vβ-nor(0.6)/nor and vβ-thal(0.6)/nor. Thus the codon 24 mutation did not appear to cause an increase in unprocessed globin RNA. Further support for this interpretation was derived from an S1 nuclease analysis with a single-stranded probe that includes most of exon 1 and intron I (33), which assays both precursor and processed mRNA. The processed mRNA-to-precursor ratio of transcripts from vβ-thal(0.6)/nor was 50% of the processed mRNA-to-precursor ratio of transcripts from vβ-nor(0.6)/nor (data not shown). The concentration of precursor in the two RNA samples appeared to be approximately equivalent.

## DISCUSSION

The substitution in codon 24 of this β-thalassemia globin gene results in activation of a splice site that is present within the coding sequence of the β-globin gene transcript. Thus this mutation differs from those that cause thalassemia by destroying the function of a normal splice site (12, 45) or by creating a new splice site within intervening sequence RNA (7, 10, 11). Orkin *et al.* (46) have recently shown that the substitution in codon 26 of the $\beta^E$ gene also leads to infrequent use of the splice site in codon 25. The mutation in codon 24 that we have characterized results in a 75% decrease in β-globin mRNA accumulation when the gene is tested in a cellular expression system, suggesting that the splice site in the coding sequence may be used 80% of the time. These observations illustrate how a subtle change in sequence may upset the exquisite precision of the splicing mechanism in generating functional globin mRNA.

Relatively little is known about the sequence requirements for accurate RNA splicing. A catalogue of the sequences found at functional splice sites (36) has led to the identification of the 5′ consensus splice sequence shown in Fig. 1. The functional splice site at the exon 1–intron I boundary of the normal human β-globin gene matches the consensus sequence in seven of nine positions, whereas the normal sequence that includes the GT in codon 25 is a six of nine match. The T-to-A substitution in codon 24 of the thalassemia gene increases the consensus match to seven of nine. This subtle change apparently converts an inactive or rarely used splice site to one that is used frequently in the $\beta^+$-thalassemia gene transcript. A small nuclear RNA ($U_1$) has also been proposed to have a role in facilitating accurate RNA splicing (47, 48). The $\beta^+$-thalassemia mutation could not affect the stability of the duplex with $U_1$ RNA because it occurs outside of the complementary segment. Our observation therefore provides direct evidence that a nucleotide not involved in the putative pairing with $U_1$ RNA may influence the function of a splice sequence.

The mechanism by which the codon 24 substitution leads to a quantitative deficiency in β-globin mRNA appears to be different than that of the G-to-A substitution in intron I that creates an alternative splice site (7, 10, 11). The aberrantly spliced RNA from the alternative splice site mutation is readily detected in bone marrow cells of both homozygous and heterozygous patients (33) and accounts for 90% of the globin RNA in tissue culture cells (10, 11). This mutation leads to an increased concentration of β-globin precursor RNA molecules from which sequences transcribed from intron I have not been removed. In contrast, the aberrantly processed RNA species arising from the codon 24 substitution is a minor species. It was readily detected in monkey kidney cells only when close proximity of the β-globin gene promoter and the SV40 enhancer element ensured a very high level of β-globin gene transcription. This RNA species could not be detected either in total bone marrow RNA (ref. 33, patient 14) or in nuclear RNA (unpublished observations). Finally, the amount of unprocessed precursor containing sequences transcribed from intron I is not increased by activation of the 5′ splice site in coding sequence. The molecular basis for the different effects of the two splice site mutations (codon 24 or intron I) is not yet understood.

The incorrectly spliced mRNA molecules arising from the splice site in codon 25 probably have 16 fewer nucleotides than normally spliced β-globin mRNA. The translation reading frame would therefore be altered, and termination of protein synthesis on this abnormal RNA would occur beyond the incorrect splice because of an in-phase termination codon. Globin mRNA molecules with premature termination codons are thought to be unstable (13–15, 17). The quantitative deficiency of total β-globin RNA transcribed from the gene with the codon 24 substitution could therefore reflect instability of the incorrectly spliced species.

Study of this $\beta^+$-thalassemia gene illustrates the usefulness of molecular cloning, DNA sequencing, and *in vitro* analysis of gene expression in characterizing naturally occurring mutations in humans. Despite intensive study, the nature of this mutation and its functional consequences in bone marrow cells were obscure until these powerful techniques were utilized. Finally, our experience illustrates how naturally occurring mutants can provide unexpected insights regarding the sequence requirements for selective and efficient RNA processing. The concept of a silent nucleotide substitution is altered by the realization that such mutations, although irrelevant with respect to the encoded amino acid, can have profound effects on RNA processing.

We are most grateful to Brian Seed and Tom Maniatis for their gift of bacterial strains, the screening miniplasmid, and detailed instructions for the miniplasmid screening procedure. We thank Anne Baur, Patricia Turner, and Austine Moulton for excellent technical assistance and express our appreciation to Linda Garza and Exa Murray for their help in the preparation of this manuscript.

1. Weatherall, D. J. & Clegg, J. B. (1981) *The Thalassemia Syndromes* (Blackwell, Oxford).
2. Weatherall, D. J. & Clegg, J. B. (1982) *Cell* **29**, 7–9.
3. Orkin, S. H., Old, J. M., Weatherall, D. J. & Nathan, D. G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2400–2404.
4. Flavell, R. A., Bernards, R., Kooter, J. M., de Boer, E., Little, P. F. R., Annison, G. & Williamson, R. (1979) *Nucleic Acids Res.* **6**, 2749–2760.
5. Orkin, S. H., Kazazian, H. H., Jr., Antonarakis, S. E., Goff, S. C., Boehm, C. D., Sexton, J. P., Waber, P. G. & Giardina, P. J. V. (1982) *Nature (London)* **296**, 627–631.

6. Poncz, M., Ballantine, M., Solowiejczyk, D., Barak, I., Schwartz, E. & Surrey, S. (1982) *J. Biol. Chem.* **257**, 5994–5996.
7. 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.
8. Westaway, D. & Williamson, R. (1981) *Nucleic Acids Res.* **9**, 1777–1788.
9. Spence, S. E., Pergolizzi, R. G., Donovan-Peluso, M., Kosche, K. A., Dobkin, C. S. & Bank, A. (1982) *Nucleic Acids Res.* **10**, 1283–1294.
10. Busslinger, M., Moschonas, N. & Flavell, R. A. (1981) *Cell* **27**, 289–298.
11. Fukumaki, Y., Ghosh, P. K., Benz, E. J., Jr., Reddy, V. B., Lebowitz, P., Forget, B. G. & Weissman, S. M. (1982) *Cell* **28**, 585–593.
12. Treisman, R., Proudfoot, N. J., Shander, M. & Maniatis, T. (1982) *Cell* **29**, 903–911.
13. Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A. & Ross, J. (1981) *Cell* **27**, 543–553.
14. Chang, J. C. & Kan, Y. W. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2886–2889.
15. Trecartin, R. F., Liebhaber, S. A., Chang, J. C., Lee, K. Y., Kan, Y. W., Furbetta, M., Angius, A. & Cao, A. (1981) *J. Clin. Invest.* **68**, 1012–1017.
16. Orkin, S. H. & Goff, S. C. (1981) *J. Biol. Chem.* **256**, 9782–9784.
17. Moschonas, N., de Boer, E., Grosveld, F. G., Dahl, H. H. M., Wright, S., Shewmaker, C. K. & Flavell, R. A. (1981) *Nucleic Acids Res.* **9**, 4391–4401.
18. Nienhuis, A. W., Turner, P. & Benz, E. J., Jr. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3960–3964.
19. Kantor, J. A., Turner, P. H. & Nienhuis, A. W. (1980) *Cell* **21**, 149–157.
20. 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.
21. Benz, E. J., Jr., Scarpa, A. L., Tonkonow, B. L., Pearson, H. A. & Ritchey, A. K. (1981) *J. Clin. Invest.* **68**, 1529–1538.
22. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
23. Kretschmer, P. J., Kaufman, R. E., Coon, H. C., Chen, M., Geist, C. E. & Nienhuis, A. W. (1980) *J. Biol. Chem.* **255**, 3204–3211.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 353–361.
25. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
27. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
28. Lawn, R. M., Efstratiadis, A., O'Connell, C. & Maniatis, T. (1980) *Cell* **21**, 647–651.
29. Humphries, R. K., Ley, T., Turner, P., Moulton, A. D. & Nienhuis, A. W. (1982) *Cell* **30**, 173–183.
30. Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. (1981) *Cell* **27**, 279–288.
31. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12**, 721–732.
32. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **7**, 1175–1193.
33. Ley, T. J., Anagnou, N. P., Pepe, G. & Nienhuis, A. W. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4775–4779.
34. Kretschmer, P. J., Coon, H. C., Davis, A., Harrison, M. & Nienhuis, A. W. (1981) *J. Biol. Chem.* **256**, 1975–1982.
35. Gorski, J., Fiori, M. & Mach, B. (1982) *J. Mol. Biol.* **154**, 537–540.
36. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
37. Jeffreys, A. J. (1979) *Cell* **18**, 1–10.
38. Tuan, D., Biro, P. A., deRiel, J. K., Lazarus, H. & Forget, B. G. (1979) *Nucleic Acids Res.* **6**, 2519–2544.
39. Antonarakis, S. E., Boehm, C. D., Giardina, P. J. V. & Kazazian, H. H., Jr. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 137–141.
40. Kan, Y. W., Lee, K. Y., Furbetta, M., Angius, A. & Cao, A. (1980) *N. Engl. J. Med.* **302**, 185–188.
41. Gluzman, Y. (1981) *Cell* **23**, 175–182.
42. Myers, R. M. & Tjian, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6491–6495.
43. Lusky, M. & Botchan, M. (1981) *Nature (London)* **293**, 79–81.
44. Banerji, J., Rusconi, S. & Schaffner, W. (1981) *Cell* **27**, 299–308.
45. Felber, B. K., Orkin, S. H. & Hamer, D. H. (1982) *Cell* **29**, 895–902.
46. Orkin, S. H., Kazazian, H. H., Jr., Antonarakis, S. E., Ostrer, H., Goff, S. C. & Sexton, J. P. (1982) *Nature (London)* **300**, 768–769.
47. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. (1980) *Nature (London)* **283**, 220–224.
48. Rogers, J. & Wall, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1877–1879.