The EMBO Journal vol.5 no.10 pp.2551-2557, 1986

β^0 Thalassemia caused by a base substitution that creates an alternative splice acceptor site in an intron

James E.Metherall, Francis S.Collins[1], Julian Pan, Sherman M.Weissman and Bernard G.Forget

Departments of Human Genetics, Internal Medicine and Molecular Biophysics and Bichemistry, Yale University School of Medicine, New Haven, CT 06510, USA

[1]Present address: Departments of Internal Medicine and Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

Communicated by A.E.Smith

A thalassemic β-globin gene cloned from a haplotype I chromosome contains a T to G transversion at position 116 of IVS1 which results in the generation of an abnormal alternative acceptor splice site. Transient expression studies revealed a 4-fold decrease in the amount of RNA produced with >99% of it being abnormally spliced despite preservation of the normal acceptor splice site at position 130. These results suggest that the mutation at IVS1 position 116 results in β^0 thalassemia. A closely related mutation at position 110 of IVS1 also generates a novel acceptor site and results in a similar decrease in total mRNA produced, but ~20% of the mRNA produced is normally spliced and thus the phenotype is that of β^+ thalassemia. These observations suggest that short range position effects may play a dramatic role in the choice of potential splice acceptor sites. We demonstrate the presence of abnormally spliced mRNA in reticulocytes of affected individuals and show the mutation at IVS1 position 116 segregating from the mutation at IVS1 position 110 in a three generation pedigree. The mutation results in the creation of a *Mae*I restriction site, as do a number of other thalassemic mutations, and we demonstrate some difficulties that may arise in the differential diagnosis of these mutations.

Key words: β-thalassemia/RNA splicing/pre-natal diagnosis

Introduction

The analysis of a large number of naturally occurring mutations in the human β-globin gene has provided a great deal of information regarding the mechanisms of normal transcription and mRNA processing. These mutations can be clinically divided into two types: β^0 thalassemia, in which there is no detectable normal β-globin polypeptide synthesized, and β^+ thalassemia, in which normal β chains are produced but in reduced amounts. The heterogeneity seen at the protein level is also present at the molecular level and naturally occurring mutations have been described that interfere with virtually every stage of protein production. A number of mutations that interfere with the removal of introns from β-globin pre-mRNA have been described and are responsible for a considerable fraction of the characterized thalassemias (for review, see Bunn and Forget, 1986; Collins and Weissman, 1984; Nienhuis *et al.*, 1984; Orkin and Kazazian, 1984).

Recently, the development of *in vitro* splicing systems has resulted in the elucidation of many biochemical characteristics of the mRNA splicing process. The sequential appearance of products in these systems suggests that the first steps of the pathway involve the cleavage of the 5′ splice site with concomitant formation of a 'lariat' structure — a structure where the 5′ end of the intron is attached by a 2′−5′ phosphodiester bond to an adenosine residue near the 3′ end of the intron (Grabowski *et al.*, 1984; Padgett *et al.*, 1984; Ruskin *et al.*, 1984). This step is followed by cleavage at the 3′ splice site with simultaneous ligation of the exons, producing a spliced RNA and an excised intron in the form of a lariat. The fidelity with which the two products of the initial cleavage event — the 5′ exon and the lariat-3′ exon — are rejoined suggests that they are maintained in close association following the cleavage event. This association appears to be accomplished by their incorporation into a multicomponent complex, which also contains pre-mRNAs, mature mRNA and the small ubiquitous ribonucleoproteins (snRNPs) containing U1 and U2 RNA (Frendewey and Keller, 1985; Grabowski *et al.*, 1985).

Although our understanding of the biochemistry of the splicing reaction has progressed rapidly, the signals within the RNA that direct the excision process have not been completely elucidated. The compilation of sequences surrounding intron boundaries has demonstrated consensus sequences at both the 5′ (donor) and the 3′ (acceptor) splice sites (Mount, 1982; Sharp, 1981; Lerner *et al.*, 1980; Rogers and Wall, 1980; Seif *et al.*, 1979). The presence of evolutionarily conserved sequences at these positions suggests that they play a role in the splicing process. This suggestion has been supported by the observation that a number of naturally occurring and *in vitro* derived mutations within the consensus regions result in splice site inactivation (Spritz and Van Santen, 1985; Wieringa and Edmonds, 1983; Treisman *et al.*, 1982, 1983; Benyajati *et al.*, 1982; Felber *et al.*, 1982; Montell *et al.*, 1982; Solnick, 1981) and by a number of observations demonstrating that these sequences are required for interaction between pre-mRNA and components of the macromolecule splicing complex (Frendewey and Keller, 1985).

It has been proposed that a 5′ to 3′ processive scanning mechanism may be involved in the choice of the 3′ acceptor site (Sharp, 1981). Experimental evidence from studies using duplicated acceptor sequences has corroborated (Lang and Spritz, 1983) as well as contradicted (Kuhne *et al.*, 1983) this hypothesis. A more recent model has proposed a limited scanning mechanism which initiates at the site of lariat formation and proceeds in a 3′ direction until an acceptor site is found (Reed and Maniatis, 1985). This proposal is based on the observation that no AG dinucleotides have been found within a region extending from 5 to 15 bp 5′ to the acceptor site (Seif *et al.*, 1979). The characterization of a β-thalassemic mutation located near the IVS1 acceptor site in the human β-globin gene (Fukamaki *et al.*, 1982; Busslinger *et al.*, 1981; Westaway and Williamson, 1981; Spritz *et al.*, 1981) also supports this notion. A G to A transition at position 110 of IVS1 results in the generation of an alternative acceptor site located 19 bp upstream of the normal acceptor site and 17 bp downstream of the site of branch point formation. Transient expression analysis has shown that the abnormal acceptor site is used preferentially over the still present normal site and the

analysis of splicing products in *in vitro* splicing systems has demonstrated that the normal branch point is utilized (Reed and Maniatis, 1985). These results suggest that the distance between the branch point and the site of cleavage is not critical and is consistent with the 5′ to 3′ scanning model in which the first AG downstream of the branch point is preferentially utilized.

Here we describe the isolation and characterization of a T to G transversion at IVS1 position 116 of the human β-globin gene which results in the generation of a splice acceptor site located only 5 bp away from the acceptor site generated by the position 110 mutation. We compare the RNA products of these genes both *in vivo* and in transient expression systems. This analysis reveals a dramatic difference in the efficiency with which the abnormal acceptor sites are utilized. Transient expression studies show that both mutations result in about a 4-fold decrease in the total amount of mRNA produced but, unlike the position 110 mutation, the position 116 mutation produces no detectable normal β-globin mRNA. These results strongly suggest that the position 116 mutation results in β^0 thalassemia, and are consistent with a 5′ to 3′ linear scanning process through the acceptor region. However, since the relative splicing efficiencies at the normal acceptor site differ in the two mutants, factors other than linear position alone must be involved in the discrimination between potential splice acceptor sites.

Results

Isolation and identification of β thalassemic gene

We have reported the identification of a point mutation associated with the Greek form of non-deletion hereditary persistence of fetal hemoglobin (ndHPFH) (Collins *et al.*, 1985). Since no homozygotes for the disorder have been identified we chose to study an individual doubly heterozygous for ndHPFH and β thalassemia. In order to unequivocally define alleles we obtained cosmid clones containing the 40-kb *Kpn*I fragment that extends from 6 kb 5′ to the Gγ gene to 3 kb 3′ to the β gene. This fragment therefore includes both of the fetal genes, Gγ and Aγ, as well as the adult genes, δ and β. Seven independent isolates were obtained. All seven gave identical restriction enzyme patterns with >20 enzymes and were derived from chromosomes of haplotype I as defined by Orkin *et al.* (1982).

Transient expression analysis was performed in order to determine the nature of the thalassemic mutation. The 5.0-kb *Bgl*II fragment containing the β-globin gene from each cosmid clone was subcloned into the expression vector LTN3b and transfected into monkey kidney (COS) cells by calcium phosphate-mediated gene transfer. Total cytoplasmic RNA was isolated after 48 h and subjected to S1 nuclease analysis using uniformly labeled M13 probes spanning the entire β-globin gene. Initial analysis demonstrated that three of the seven clones produced RNA with abnormal sized exon 2 (data not shown). Subsequent analysis using a 5′ end-labeled *Bam*HI fragment localized this abnormality to the IVS1 acceptor site (Figure 1). Normally spliced RNA results in the protection of a 209-nucleotide fragment, as shown in lane 8. Four of the β gene isolates (lanes 1−3, 5) produced mRNA which protected the normal sized fragment, while three others protected a fragment which was slightly longer (lanes 4, 6−7). DNA sequence analysis revealed a single base change at position 116 of IVS1 (data not shown). The T to G transversion results in the generation of an AG dinucleotide only 14 bp upstream of the invariant AG dinucleotide found at the normal acceptor site (Table I). Splicing at this novel AG results in the inclusion of IVS1 sequences in mature mRNA producing a frameshift within exon 2 where a terminator codon is reached 34 amino acids downstream from the abnormal splice.

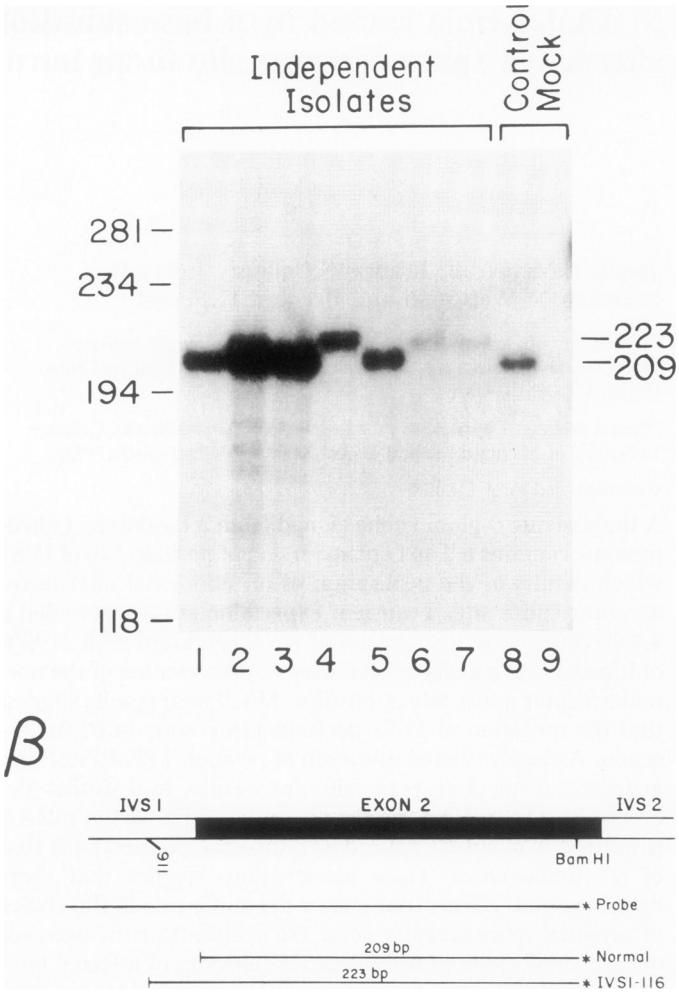

Fig. 1. Identification of thalassemic globin genes using S1 nuclease analysis. Cytoplasmic RNA was isolated from COS cells 36−48 h following transfection with β-globin genes cloned in the expression vector LTN3b. S1 nuclease analysis was performed using an end-labeled 1.8-kb *Bam*HI fragment. The resulting protected fragment extends from the *Bam*HI site within exon 2 to the IVS1 acceptor site. In normal RNA this coresponds to a 209-nucleotide fragment, as seen in **lane 8**. Seven independent clones were isolated from an individual doubly heterozygous for β thalassemia and ndHPFH. Four of the clones **(lanes 1−3,5)** demonstrate a normally sized protection product while three others **(lanes 4,6−7)** show a larger product.

IVS1 position 116 mutation results in β^0 thalassemia

During the identification of the IVS1 position 116 mutation we were surprised by the efficiency with which the abnormal acceptor site was utilized. No detectable normal RNA was produced as is demonstrated most clearly by lane 4 in Figure 1. We then performed a more quantitative analysis of the RNA produced from both the IVS1 position 110 and 116 mutations by transient expression analysis in HeLa cells, using co-transfected α- or γ-globin genes to control for transfection efficiency.

Transient expression analysis of β-globin genes in HeLa cells has been shown to accurately reflect the *in vivo* characteristics of a number of thalassemic splicing mutations. A mutation at position 5 of IVS1 results in the partial inactivation of the donor site in this system while a mutation at IVS1 position 6 results in a less dramatic decrease in the level of normal RNA produced

Table I. Mutations generating novel acceptor splice sites in human β-globin IVS1

g Leu Leu Val c t c t c t c t g c c t a t t g g t c t a t t t t c c c a c c c t t a g GCTGCTGGTGG	Normal
g Leu Phe Ser His Pro Ter c t c t c t c t g c c t a t t a g TCTATTTTCCCACCCTTAGGCTGCTGGTGG *	IVS1−110
g Phe Pro Pro Leu Gly Cys Trp Trp c t c t c t c t g c c t a t t g g t c t a g TTTCCCACCCTTAGGCTGCTGGTGG *	IVS1−116

The DNA sequence surrounding the IVS1 splice acceptor site is shown. Intron sequences are shown in lower case and the amino acid translations are given above the DNA sequence. The invariant AG dinucleotide at the normal acceptor site is underlined and asterisks show the location of the mutations at positions 110 and 116.

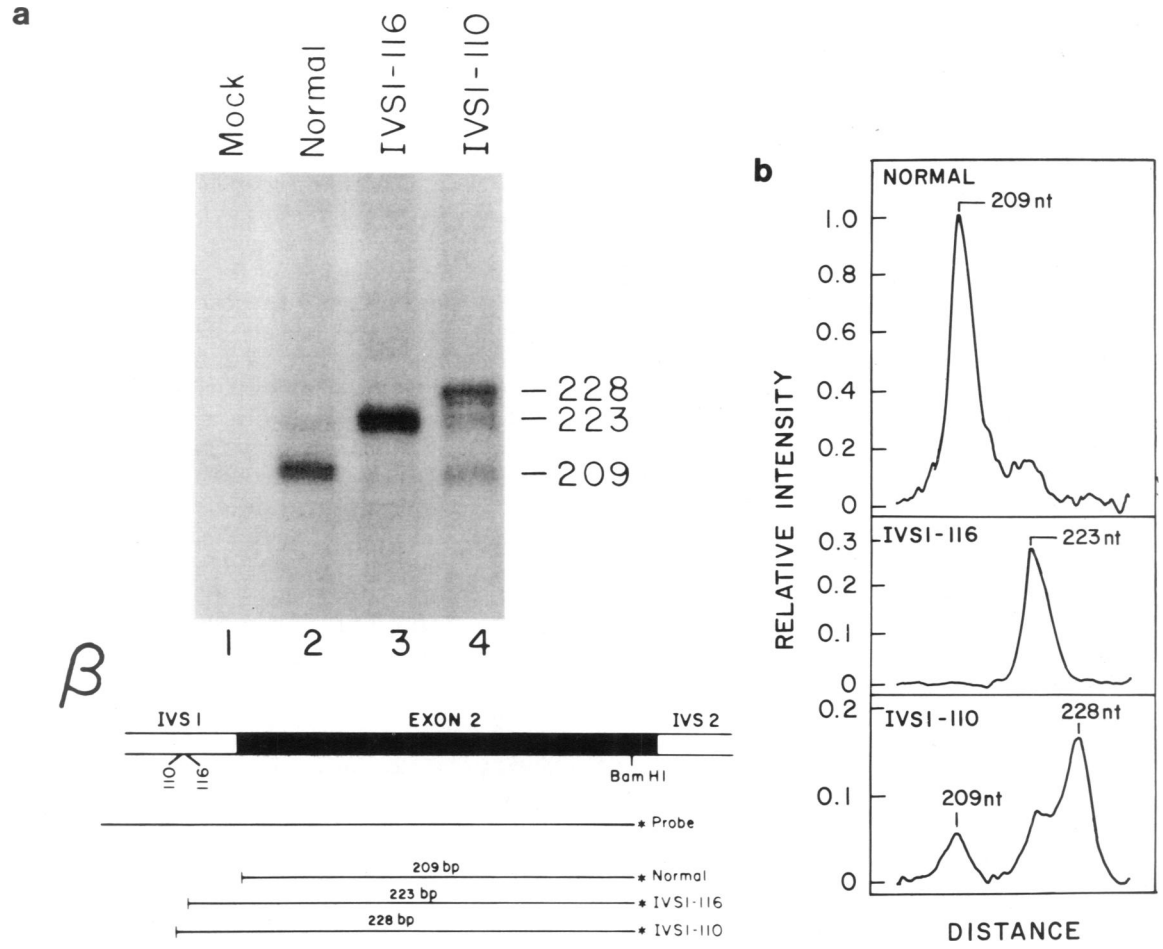

Fig. 2. Quantitation of RNA produced from IVS1 acceptor site mutations. **a** shows a typical S1 analysis of RNA isolated from HeLa cells transiently expressing β-globin genes. **Lane 1** shows a mock transfection control. The fragments protected by the normal **(lane 2)** and mutant genes **(lanes 3,4)** are diagrammed below the figure. **b** shows composite densitometric tracings of this and other experiments following normalization for transfection efficiency using a co-transfected control gene as well as a point-by-point subtraction of background as represented in the mock-transfected controls. Note the differences in scale between the mutant and normal genes.

(Treisman *et al.*, 1983). These observations correlate well with clinical severity. Although both mutations result in β^+ thalassemia, the position 5 mutation produces a more pronounced deficiency of β chain synthesis than does the position 6 mutation.

β-Globin genes were co-transfected into HeLa cells along with either a normal α-globin gene or a normal γ-globin gene in order to quantitate transfection efficiency. Total cytoplasmic RNA was subjected to S1 nuclease analysis using end-labeled DNA probes. An example of this is shown in Figure 2a, where an end-labeled *Bam*HI fragment was used as a probe. Lane 2 shows the 209-nucleotide fragment protected by normally spliced RNA. Lanes 3 and 4 show the fragments protected by the transcripts of the IVS1 position 116 and 110 genes, respectively. The position 110 mutant gene clearly produces some normal β mRNA (lane 4)

Table II. Summary of splice site efficiencies

			Relative efficiency
Consensus		Y Y Y Y Y Y Y Y Y Y Y N Y A G / G	
Normal		T T T T C C C a C C C T T A G / G	1.00
IVS1-110	Mutant	C T C T C T g C C T a T T A G / t	0.19
	Normal	T T T T C C C a C C C T T A G / G	0.05
IVS1-116	Mutant	T g C C T a T T g g T C T A G / t	0.27
	Normal	T T T T C C C a C C C T T A G / G	0.00

The splice sites generated by the position 116 and 110 mutations are shown, along with their relative efficiencies. Splice site sequences are shown below the consensus sequence (Mount, 1982) and nucleotides not matching the consensus are shown in lower case.

while the position 116 mutant gene produces little or no normal β mRNA (lane 3). Prolonged exposure revealed the presence of a diffuse band at the position expected for the normal β mRNA, but this band was also present in the mock-transfected control (lane 1). In order to more accurately determine whether a component of this band could be attributed to normally spliced RNA, densitometric tracings were obtained from a number of such experiments and digitized to allow a point-by-point subtraction of background. A composite of these tracings is shown in Figure 2b. The upper panel shows a composite tracing for the normal gene and the lower two panels show similar tracings for the two mutant genes. The middle panel shows quite clearly that the IVS1 position 116 gene produces no detectable normally spliced β-globin mRNA in this assay. Dilution analysis of a number of RNA samples was performed to determine the limits of detection of this assay (data not shown). We conclude from these analyses that >99% of the RNA produced from this gene is abnormally spliced.

A summary of the relative efficiencies of the various splice sites is presented in Table II. Both mutations result in an ~4-fold decrease in total mRNA production. Virtually no normal mRNA is produced from the IVS1 position 116 mutation while >20% of the RNA produced from the IVS1 position 110 mutation is normally spliced. We believe that the lack of normal RNA with the position 116 mutation is so severe that it results in β^0 thalassemia. Verification of the phenotype awaits the identification of a homozygous individual or an individual carrying the mutation in *trans* to a known β^0 thalassemic gene.

In vivo characterization of IVS1 position 116 mutation

The IVS1 position 116 mutation appears to be a relatively infrequent mutation in the Mediterranean population. A survey of 63 haplotype I thalassemic genes using oligonucleotide hybridization and direct restriction enzyme analysis revealed only two cases where the mutation was not attributable to a previously described mutation (Kazazian *et al.*, 1984) and therefore could possibly be due to the IVS1 position 116 mutation. A survey of S1 nuclease analysis of the IVS1−exon 2 junction from a large number of thalassemic patients revealed only a single individual with an abnormal pattern that could be attributed to the position 116 mutation (Benz *et al.*, 1982). Reticulocyte RNA was prepared from this individual as well as from a number of his first degree relatives and from the individual from whom the IVS1 position 116 gene was originally isolated. The RNA was then subjected to S1 nuclease analysis as shown in Figure 3. The fragments protected by reticulocyte RNA from a normal individual, from an individual homozygous for the IVS1 position 110 mutation and from the heterozygote for the IVS1 position 116 mutation are

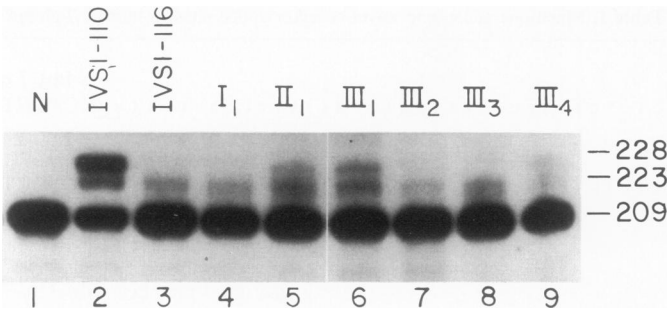

Fig. 3. S1 nuclease analysis of reticulocyte RNA. S1 nuclease analysis was performed using the *Bam*HI-labeled probe described in Figure 2a on RNA from a normal individual **(lane 1)**, an individual believed to be homozygous for the IVS1 position 110 mutation **(lane 2)**, the individual from whom the IVS1 position 116 mutation was cloned **(lane 3)** and from an individual heterozygous for the two mutations **(lane 5)**, as demonstrated by the segregation of the two mutations in his father **(lane 4)** and his children **(lanes 6–9)**.

shown in lanes 1, 2 and 3, respectively. The position 116 mutation is clearly distinguishable from the position 110 mutation as demonstrated by the difference between lanes 2 and 3. The minor band of intermediate mobility seen in the lane with the position 110 mutation seems to be due to overdigestion of the abnormally spliced product, since lowering the reaction temperature reduces the amount of the minor band present (data not shown). Lower reaction temperatures were not routinely used because an S1 artifact band of ~220 nucleotides is seen in normal RNA (Busslinger *et al.*, 1981), making it difficult to identify the abnormal product of the position 116 mutation.

The individual originally identified by Benz *et al.* (1982) as having an abnormal splicing pattern at the IVS1−exon 2 junction (lane 5) appears to be heterozygous for the two acceptor site mutations as demonstrated by the segregation of the mutations among his first degree relatives (lanes 4, 6–9). He has inherited the IVS1 position 116 mutation from his father (lane 4) and has transmitted the IVS1 position 116 mutation to two of his children (lanes 7, 8) and the IVS1 position 110 mutation to his other two children (lanes 6, 9). The normally spliced mRNA produced from the position 110 gene has been shown to selectively accumulate in reticulocytes of affected individuals (Fukamaki *et al.*, 1982), presumably due to the instability of the abnormally spliced mRNA. This phenomenon probably explains the high levels of normally processed β mRNA relative to mutant β mRNA found in the reticulocytes of the β^0/β^+ heterozygote. This family has α and β thalassemic traits co-segregating within the pedigree and was originally reported to demonstrate the clinically advantageous interaction between α and β thalassemic alleles due to the normalization of α:β globin chain synthetic ratios (Kan and Nathan, 1970).

MaeI recognition of β thalassemia mutations

The restriction enzyme *Mae*I has recently been identified (Schmid *et al.*, 1984). It recognizese the tetrameric sequence CTAG and allows the direct detection of the IVS1 position 116 mutation. This enzyme has also been shown to be useful in the diagnosis of the most common form of β^0 thalassemia in the Mediterranean population (Thein *et al.*, 1985), specifically a mutation at codon 39 which results in the generation of a terminator codon. In addition *Mae*I has been postulated to recognize a second termination mutation at codon 17 (Thein *et al.*, 1985). All three of these mutations reside in the same region of the β-globin gene. In order

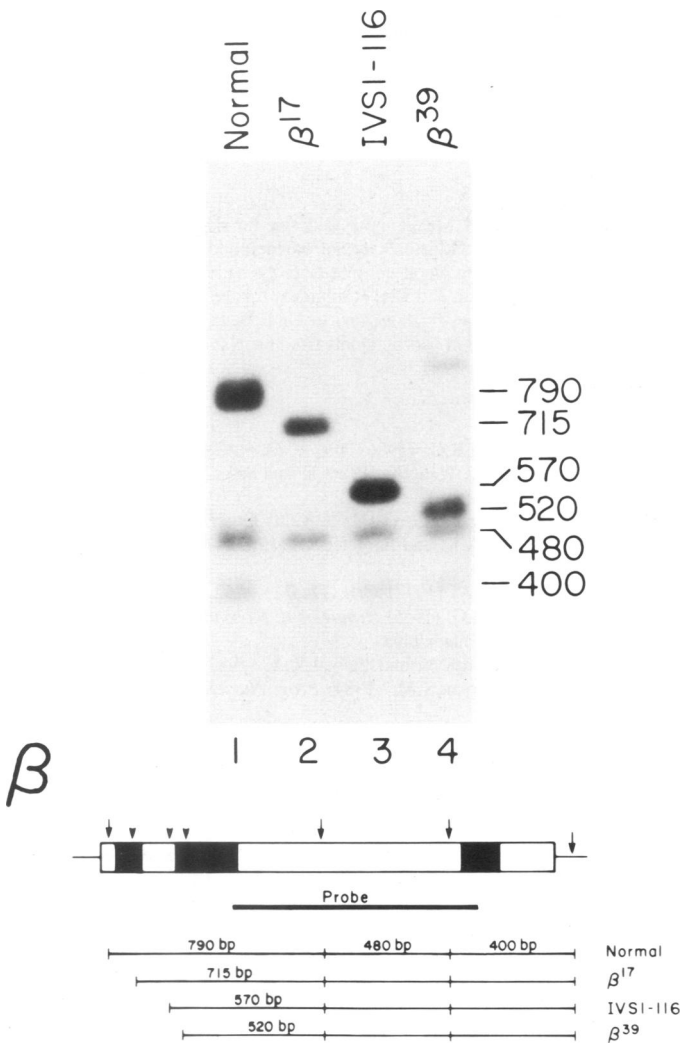

Fig. 4. DNA blot analysis of thalassemic mutations recognized by *Mae*I. Plasmid DNA was blotted using the method of Southern (1975) following digestion with *Mae*I and was probed using an IVS2 probe extending from the *Bam*HI site through the *Eco*RI site. The probe hybridizes to three bands in each case. In the normal **(lane 1)** the bands are 400 bp, 480 bp and 790 bp. The thalassemic mutations shorten the 790-bp fragment to 715 bp in β^{17} **(lane 2)**, to 570 bp in IVS1 position 116 **(lane 3)** and to 520 bp in β^{39} **(lane 4)**.

to determine whether these mutations could be differentially diagnosed, plasmid DNA of cloned normal and thalassemic β-globin genes was digested with *Mae*I, resolved by electrophoresis on a 2% agarose gel and hybridized to an IVS2 probe using the method of Southern (1975). The results of this analysis are shown in Figure 4. Lane 1 shows the normal restriction pattern, while lanes 2–4 show the patterns produced by the various thalassemic mutations. All three thalassemic mutations are distinguishable, mainly due to the existence of reference bands in each lane which result from hybridization of the probe to adjacent *Mae*I fragments. We routinely experienced difficulty in digesting the *Mae*I site at the β^{39} position as demonstrated by the presence of a partial digest in lane 4.

Discussion

Here we describe the isolation and characterization of a novel mutation resulting in β thalassemia. This mutation is interesting for a number of reasons. It is the second thalassemic point mutation to be described that occurs near the 3′ end of IVS1 and allows the critical evaluation of a number of characteristics regarding the mechanism of splicing. It also reveals a potential problem in the differential diagnosis of a number of thalassemic mutations using DNA blotting techniques.

The position 116 mutation results in the production of little or no normally spliced mRNA. Previous studies of a closely related mutation at IVS1 position 110 implicated a 5′ to 3′ scanning process in the choice of splice acceptor site. The results presented here support this hypothesis by demonstrating that a second mutation which creates an AG dinucleotide near the 3′ splice site also results in the preferential use of the more 5′ acceptor site.

The IVS1 position 116 mutation differs from the IVS1 position 110 mutation in that it does not produce any normally spliced β-globin mRNA. It is possible that this difference is due to the creation of a better splice site in the case of the IVS1 position 116 mutation, but this is unlikely since it diverges further from the consensus sequence than does the IVS1 position 110 splice site. A second possibility is that the mutation at position 116 has two effects: one being the creation of a new acceptor site and the other being the disruption of the normal acceptor site. The mutation shortens the polypyrimidine tract of the normal site from 11 to 10 bp. The significance of this difference is not clear. Conservation at this position is demonstrated by the presence of pyrimidines in 70% of known splice acceptor sites, but >15% of the splice sites contain G residues at this position, including IVS1 of the human ε-globin gene (Mount, 1982). Clearly the presence of a G at this position is compatible with efficient splicing. A third possibility to explain the presence of normally spliced RNA with the IVS1 position 110 mutation and not with the IVS1 position 116 mutation is that the position 110 mutation creates a splice acceptor site which is located too close to the branch point to be used exclusively, possibly due to steric interference between the factors that bind at the branch point and the factors which bind at the acceptor site.

An interesting set of observations was made in the study of deletions involving the IVS2 acceptor site of the rabbit β-globin gene (Wieringa *et al.*, 1984), where it was concluded that the efficiency of 3′ cleavage is determined by incremental contributions of the polypyrimidine tract. A factor that complicated this analysis was the presence of a cryptic splice site in the sequences that were juxtaposed to the normal acceptor. The larger the deletion, the closer the cryptic site was brought to the normal site. When the cryptic site was >24 bp away from the normal site, the normal acceptor was used exclusively; but when the cryptic site was brought to within 21 bp, both normal and cryptic sites were used equally and when the cryptic site was brought closer than 21 bp the cryptic site was used exclusively. These observations are consistent with the findings presented here. The IVS1 position 110 mutation creates a cryptic site which is located 19 bp upstream of the normal site and results in utilization of both sites, while the IVS1 position 116 mutation creates an acceptor site located only 14 bp upstream of the normal site and results in exclusive use of the abnormal site. Further studies involving both mutant and normal splice sites are needed to determine the significance of these observations and should provide valuable information regarding the mechanism of splicing.

The second interesting aspect of the IVS1 position 116 mutation is that it reveals some problems that can arise using DNA-based diagnosis. Our understanding of the etiology of hemoglobinopathies has progressed to the point where individual molecular lesions can be assayed using DNA hybridization techniques. This information has led to the use of oligonucleotide

hybridization and DNA blot analysis in pre-natal diagnosis. The ability to accurately diagnose these mutations is critical in the establishment of an effective pre-natal program. The need for accuracy has recently been demonstrated by the potential pitfall in diagnosing a thalassemic mutation at IVS1 position 6. The diagnosis, based on DNA blotting using the restriction enzyme *Sfa*NI, could be inaccurate if one fails to recognize the presence of a polymorphic *Sfa*NI site located at codon 2 in some β-globin genes (Atweh and Forget, 1986).

The demonstration that the IVS1 position 116 mutation creates an *Mae*I site raises the question as to whether this mutation can be distinguished from other thalassemic mutations which also create *Mae*I sites. One of the most common forms of β^0 thalassemia results from a mutation which creates a translational terminator at codon 39. This and a similar mutation at codon 17 also generates *Mae*I sites (Thein *et al.*, 1985), which are located within the same wild-type *Mae*I fragment. The ability to distinguish these mutations is demonstrated by the blot in Figure 6. The difference between the normal and each of the mutants is clearly discernible, but the ability to distinguish between the β^{39} and the IVS1 position 116 mutations relies heavily on the presence of internal reference bands. These reference bands result from the hybridization of the probe to surrounding *Mae*I fragments and without them differences could easily be attributed to anomalous gel migration. We also encountered specific resistance of the β^{39} site to *Mae*I digestion using a number of plasmid DNA preparations. This problem would probably be compounded in the case of genomic DNA digests. Such specific partials may easily result in the inaccurate diagnosis of β^{39} homozygotes as heterozygotes.

Materials and methods

Clones

Cosmid clones containing the 40-kb *Kpn*I fragment bearing the non-α-globin gene cluster were obtained as previously described (Collins *et al.*, 1984). The 5.0-kb fragments containing the β-globin genes from seven independent cosmids were cloned into the SV40 enhancer-containing expression vector LTN3b (Humphries *et al.*, 1982). The isolation of the IVS1 position 110 gene and the β^{39} gene in the expression vector LTN3b has been previously described (Takeshita *et al.*, 1984). The β^{17} gene was created by site-directed mutagenesis and was the kind gift of Susan Baserga.

Sequencing

The 5.0-kb *Bgl*II fragments containing the β-globin gene were isolated and digested to completion with the restriction enzymes *Hae*III and *Rsa*I. These fragments were cloned into the *Sma*I site of the phage M13 mp8 and sequenced using the dideoxy sequencing method of Sanger *et al.* (1977). T-tracks were used to determine the origin of the M13 clones and clones containing the *Hae*III fragment extending from the 3′ end of exon 1 through 130 bp of exon 2 were sequenced in their entirety.

Reticulocyte RNA isolation

Reticulocyte RNA was prepared using the detergent, tri-iso-propylnaphthalene sulfonic acid as previously described (Benz and Forget, 1971).

Tissue culture and transient expression

Monkey kidney (COS) cells and human HeLa cells were maintained in Dulbecco's modified Eagle's medium containing 10% fetal calf serum, 10 units/ml penicillin and 10 μg/ml streptomycin. Transfection of monkey kidney (COS) cells was performed by calcium phosphate-mediated gene transfer (Wigler *et al.*, 1979). Typically 4–6 100 mm dishes were used per transfection. Total cytoplasmic RNA was prepared in the presence of vanadyl-ribonucleosides (Favaloro *et al.*, 1980).

S1 nuclease analysis

S1 nuclease analysis (Berk and Sharp, 1977) was performed using the probes described in the figure legends. Probes were either end-labeled with polynucleotide kinase (Maxam and Gilbert, 1980) or cloned in M13 phage and uniformly labeled using the method of Sanger *et al.* (1977) in the absence of dideoxynucleotides. Autoradiographic signals were scanned using a double-beam microdensitometer (Joyce, Loebl and Co., Inc.) and digitized into an IBM personal computer using a Hewlett-Packard 7470 Plotter.

DNA blot hybridization

Cloned DNA was digested with the restriction enzyme *Mae*I, resolved on a 2% agarose gel and transferred to nitrocellulose using the method of Southern (1975). A 917-bp *Bam*HI–*Eco*RI fragment encompassing IVS2 of the β-globin gene was purified, labeled by nick translation and used as a hybridization probe.

Acknowledgements

We thank M.Newman, K.Miceli and F.Gillespie for skilled technical assistance and Drs N.Anagnou, D.G.Nathan, K.Sofroniadou and G.Stamatoyannopoulos for assistance in providing blood samples from the patients studied. We are also grateful to Dr J.Eisenstadt and Dr W.Summers for the use of equipment, to S. Baserga for supplying the β^{17} clone and to Dr E.Benz for helpful discussions. This work was supported in part by grants from the National Institutes of Health.

References

Atweh,G.F. and Forget,B.G. (1986) *Am. J. Hum. Genet.*, **38**, 855–859.
Benyajati,C., Place,A.R., Wang,N., Pentz,E. and Sofer,W. (1982) *Nucleic Acids Res.*, **10**, 7261–7272.
Benz,E.J., Jr. and Forget,B.G. (1971) *J. Clin. Invest.*, **50**, 2755–2760.
Benz,E.J., Jr., Scarpa,A.L. and Forget,B.G. (1982) *Trans. Assoc. Am. Phys.*, **95**, 325–333.
Berk,A.J. and Sharp,P.A. (1977) *Cell*, **12**, 721–732.
Bunn,H.F. and Forget,B.G. (1986) *Hemoglobin: Molecular, Genetic and Clinical Aspects*. Saunders, Philadelphia.
Busslinger,M., Moschonas,N. and Flavell,R.A. (1981) *Cell*, **27**, 289–298.
Collins,F.S. and Weissman,S.M. (1984) *Prog. Nucleic Acid Res. Mol. Biol.*, **31**, 315–458.
Collins,F.S., Stoeckert,C.J., Serjeant,G.R., Forget,B.G. and Weissman,S.M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4894–4898.
Collins,F.S., Metherall,J.E., Yamakawa,M., Pan,J., Weissman,S.M. and Forget, B.G. (1985) *Nature*, **313**, 325–326.
Favaloro,J., Treisman,R. and Kamen,R. (1980) *Methods Enzymol.*, **65**, 718–749.
Felber,B.K., Orkin,S.H. and Hamer,D.H. (1982) *Cell*, **29**, 895–902.
Frendewey,D. and Keller,W. (1985) *Cell*, **42**, 355–367.
Fukamaki,Y., Ghosh,P.K., Benz,E.J., Reddy,V.B., Lebowitz,P., Forget,B.G. and Weissman,S.M. (1982) *Cell*, **28**, 585–593.
Grabowski,P.J., Padgett,R.A. and Sharp,P.A. (1984) *Cell*, **37**, 415–427.
Grabowski,P.J., Seiler,S.R. and Sharp,P.A. (1985) *Cell*, **42**, 345–353.
Humphries,R.K., Ley,T., Turner,P., Moulton,A.D. and Nienhuis,A.W. (1982) *Cell*, **30**, 173–183.
Kan,Y.W. and Nathan,D.G. (1970) *J. Clin. Invest.*, **49**, 635–642.
Kazazian,H.H., Orkin,S.H., Markham,A.F., Chapman,C.R., Youssoufian,H. and Waber,P.C. (1984) *Nature*, **310**, 152–154.
Kuhne,T., Wieringa,B., Reiser,J. and Weissmann,C. (1983) *EMBO J.*, **2**, 727–733.
Lang,R.M. and Spritz,R.A. (1983) *Science*, **220**, 1351–1355.
Lerner,M., Boyle,J., Mount,S., Wolin,S. and Steitz,J. (1980) *Nature*, **283**, 220–224.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 497–559.
Montell,C., Fisher,E.F., Caruthers,M.H. and Berk,A.J. (1982) *Nature*, **295**, 380–384.
Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.
Nienhuis,A.W., Anagnou,N.P. and Ley,T.J. (1984) *Blood*, **63**, 738–758.
Orkin,S.H. and Kazazian,H.H. (1984) *Annu. Rev. Genet.*, **18**, 131–171.
Orkin,S.H., Kazazian,H.H., Antonaarakis,S.E., Goff,S.C., Boehm,C.D., Sexton, J.P., Waber,P.G. and Giardina,P.J.V. (1982) *Nature*, **296**, 627–631.
Padgett,R.A., Konarska,M.M., Grabowski,P.J., Hardy,S.F. and Sharp,P.A. (1984) *Science*, **225**, 898–903.
Reed,R. and Maniatis,T. (1985) *Cell*, **41**, 95–105.
Rogers,J. and Wall,R. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1877–1879.
Ruskin,B., Krainer,A.R., Maniatis,T. and Green,M.R. (1984) *Cell*, **38**, 317–331.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schmid,K., Thomm,M., Laminet,A., Laue,F.G., Kessler,C., Stetter,K.O. and Schmitt,R. (1984) *Nucleic Acids Res.*, **12**, 2619–2628.
Seif,I., Khoury,G. and Dhar,R. (1979) *Nucleic Acids Res.*, **6**, 3387–3389.
Sharp,P. (1981) *Cell*, **23**, 643–646.
Solnick,D. (1981) *Nature*, **291**, 508–510.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503.
Spritz,R.A. and Van Santen,V.L. (1985) *Ann. N.Y. Acad. Sci.*, **445**, 10–19.
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. USA*, **78**, 2455–2459.
Takeshita,K., Forget,B.G., Scarpa,A. and Benz,E.J. (1984) *Blood*, **64**, 13–22.

Thein,S.L., Wainscoat,J.S., Lynch,J.R., Weatherall,D.J., Sampietro,M. and Fiorelli,G. (1985) *Lancet*, **i**, 1095.
Treisman,R., Proudfoot,N.J., Shander,M. and Maniatis,T. (1982) *Cell*, **29**, 903–911.
Treisman,R., Orkin,S.H. and Maniatis,T. (1983) *Nature*, **302**, 591–596.
Westaway,D. and Williamson,R. (1981) *Nucleic Acids Res.*, **9**, 1777–1788.
Wieringa,J.C. and Edmonds,M. (1983) *Nature*, **301**, 38–43.
Wieringa,J.C., Hofer,E. and Weissmann,C. (1984) *Cell*, **37**, 915–925.
Wigler,M., Sweet,R., Sim,G.K., Wold,B., Pellicer,A., Lacy,E., Maniatis,T., Silverstein,S. and Axel,R. (1979) *Cell*, **16**, 777–785.

Received on 22 May 1986; revised on 15 July 1986