# Molecular basis of an apolipoprotein[a] null allele: a splice site mutation is associated with deletion of a single exon

**Laura A. Cox, Catherine Jett, and James E. Hixson[1]**

Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78245

**Abstract** **Apolipoprotein[a] (apo[a]), a unique component of atherogenic lipoprotein[a], is highly polymorphic in human and nonhuman primates. Null alleles, producing no detectable circulating Lp[a] or apo[a] isoforms, are found at high frequencies. The molecular basis of null alleles is not yet known. In baboons, approximately two-thirds of null alleles do not produce detectable hepatic transcripts (transcript negative nulls), and one-third of null alleles produce normal amounts of apo[a] transcripts (transcript positive nulls). We have cloned apo[a] cDNA from a baboon carrying a transcript positive null allele defective in secretion from primary hepatocytes. Compared with wild-type cDNA, the null allele contained an in-frame 47 amino acid deletion in the protease domain corresponding to one exon of the apo[a] gene. The null allele contains an A→T substitution in the third nucleotide position of the intron downstream of the deleted exon which alters the donor splice site consensus sequence. Thus, this null is likely due to a mutation that prevents normal mRNA splicing, yielding a shortened protein that may be defective in intramolecular interactions required for normal processing and secretion of apo[a]. This is the first report of a molecular basis for apo[a] null alleles.—**Cox, L. A., C. Jett, and J. E. Hixson. **Molecular basis of an apolipoprotein[a] null allele: a splice site mutation is associated with deletion of a single exon. *J. Lipid Res.* 1998.** 39: **1319–1326.**

**Supplementary key words** protein secretion • plasminogen • lipid metabolism • transcript processing

Lipoprotein[a] (Lp[a]) is a low density lipoprotein (LDL) particle that contains apo[a] covalently linked to apoB. Lp[a] has been found to accumulate in atherosclerotic plaques (1, 2), and serum levels of Lp[a] correlate with risk of cardiovascular disease independently of such factors as LDL levels, HDL levels, and levels of apolipoproteins (3). In addition, Lp[a] is considered an independent risk factor for myocardial infarction (3, 4). Serum Lp[a] concentration is a heritable trait, and is not strongly influenced by age or diet. The apo[a] locus itself accounts for over 90% of the variation in Lp[a] levels in human populations (5, 6).

Human apo[a] consists of structural domains that are also present in plasminogen, including a protease domain and Kringles IV and V (7). However, the protease domain in apo[a] has been inactivated by amino acid substitutions, and Kringle IV has been highly amplified to form large numbers of tandemly oriented repeats. Apo[a] is highly polymorphic with respect to size, varying between 400 and 700 kilodaltons, due to differences in numbers of Kringle IV repeats (8). In addition to the marked size variation of apo[a], both humans and baboons carry high frequencies of null alleles with little or no detectable apo[a] isoforms or serum Lp[a]. In humans, null allele frequencies vary between 0.29 and 0.45 for different populations. In baboons, null allele frequencies vary between 0.22 and 0.46 (9). Little is known about the molecular basis of apo[a] null alleles or the physiological consequences of carrying apo[a] null alleles. In baboons, we have shown that approximately two-thirds of null alleles do not produce hepatic apo[a] transcripts that are detectable by Northern blot analysis (called transcript negative nulls) (10). Approximately one-third of null alleles produce normal levels of hepatic apo[a] transcripts (transcript positive nulls).

In previous studies, White et al. (11) used cultured primary baboon hepatocytes to study intracellular processing and secretion of apo[a]. These studies have shown that larger apo[a] alleles have slower rates of intracellular processing, offering the first mechanistic explanation for the inverse correlation between apo[a] size and serum Lp[a] concentration observed in human and baboon populations. White, Guerra, and Lanford (12) also studied primary hepatocytes carrying apo[a] transcript positive null alleles. Their results show that apo[a] transcripts are translated, but the protein is not released from the endo-

Abbreviations: apo[a], apolipoprotein [a]; Lp[a], lipoprotein [a]; UTR, untranslated region; RT-PCR, reverse transcriptase-mediated polymerase chain reaction; PCR, polymerase chain reaction.

[1] To whom correspondence should be addressed.

plasmic reticulum during processing. These results suggest that transcript positive null alleles contain changes within the apo[a] sequence that impair normal apo[a] processing and secretion. In this study, we investigate the molecular bases for a transcript positive null allele that was previously shown to be defective in apo[a] secretion.

## EXPERIMENTAL PROCEDURES

### Sample selection

Archived liver and DNA samples were obtained from a baboon that carried a transcript positive null apo[a] allele, and from an unrelated baboon that was homozygous for a wild-type allele. Studies of primary hepatocytes from the baboon carrying the transcript positive null allele showed normal transcription and translation, but defective secretion of apo[a] protein into the medium (12). Studies of primary hepatocytes from the baboon homozygous for the wild-type allele showed normal transcription, translation, and secretion (12). We also used archived liver and DNA samples from other family members of the baboon carrying the apo[a] null allele for analysis of cDNA and genomic DNA sequences. Additional baboons that were screened for the presence of the splice site mutation were chosen based on Lp[a] phenotype and Northern blot analyses (10).

### Cloning and sequencing of apo[a] cDNAs

A cDNA library was constructed from hepatic mRNA from the baboon carrying an apo[a] transcript positive null allele. First strand cDNA synthesis used reverse transcriptase (Superscript, Gibco BRL) and a mixture of oligo d(T) and random primers. After addition of synthetic EcoRI linkers, cDNAs were ligated into lambda ZAP express vectors (Stratagene Inc.). The cDNA library was screened by hybridization with a human apo[a] probe containing only Kringle IV (8), and with a plasminogen cDNA probe containing only Kringles I–III (13) to identify cross-reacting plasminogen cDNA clones. Approximately $1 \times 10^7$ plaques were screened, yielding 116 clones that were positive for apo[a] and negative for plasminogen. The apo[a] cDNA clones containing the largest inserts were sequenced with an automated DNA sequencer (Perkin Elmer Inc.) using the dye terminator method and primers from flanking vector and internal sequences. Three different apo[a] clones were sequenced on both strands using overlapping primers in each direction.

Apo[a]-specific primers were designed from the baboon cDNA sequences, and used for reverse transcriptase mediated PCR (RT-PCR) to amplify and clone cDNA for the wild type apo[a] allele. RT-PCR used a forward primer located at the junction of the last Kringle IV and the protease domain (5′-AGGCTCGTTCTGGACAAGCATCCT-3′), and a reverse primer located in the 3′ untranslated region (3′ UTR) (5′-GTCACCTTATTGGAGAAACCAGC-3′). PCR conditions included initial denaturation (94°C for 4 min), followed by 30 cycles of denaturation (94° for 1 min) and annealing/extension (60° for 1 min), and a final extension (72° for 7 min). RT-PCR products were inserted into pCR2.1 vectors (Invitrogen) for automated sequencing. In addition, RT-PCR products were directly sequenced using the automated DNA sequencer. All mRNA samples were subjected to a minimum of two RT-PCR assays. Three independent clones were sequenced from each RT-PCR assay. All RT-PCR clones were sequenced on both strands with overlapping primers. Nucleotide sequences were aligned using Sequencher software (Gene Codes Inc.), and percent identities for nucleotide and predicted amino acid sequences were calculated using BLAST/Genbank programs (NCBI).

### RNA isolation and analyses

Messenger RNA was extracted directly from frozen baboon liver tissue using mRNA FastTrack Kit (Invitrogen). RT-PCR was used to analyze the length of apo[a] transcripts in liver samples from baboons carrying various apo[a] alleles. First strand cDNA synthesis used 100–200 ng hepatic mRNA with a reverse primer (5′-GTCACCTTATTGGAGAAACCAGC-3′) in the 3′ untranslated region (3′ UTR) and 20 U/μl reverse transcriptase under reaction conditions described by the supplier (Gibco BRL). An aliquot (10%) of the first strand cDNA reaction was amplified using the same reverse primer and a forward primer (5′-AGG CTCGTTCTGGACAAGCATCCT-3′) from the Kringle IV/protease domain junction. PCR conditions included initial denaturation (94° for 4 min), followed by 30 cycles of denaturation (94° for 1 min) and annealing/extension (60° for 1 min), and a final extension (72° for 7 min). The PCR products were subjected to electrophoresis on agarose gels (1.0%) and visualized by staining with ethidium bromide.

### Cloning and sequencing of apo[a] genomic sequences

To examine genomic sequences of the wild-type baboon apo[a] and plasminogen genes, lymphocyte DNA was amplified with primers from reported human gene sequences (14, 15), and PCR products were cloned for nucleotide sequencing. Two overlapping fragments were amplified including a fragment containing exons E, F, and G, and intervening introns (forward primer 5′-CTCAGGATCCATCCTCTTCATTTGATTGTGGGAAG-3′, reverse primer 5′-TGCTAAGCTTTAGCAAGGCAATATCTGCTTG-3′); and a second fragment containing part of exon G, exon H, and the intervening introns (forward primer 5′-CTGGGAAT TCACCAAGAAGTGAACCTCGAATCTCA-3′, reverse primer 5′-AAGGAAGCTTGGGTTTCTCCCCAGCCAGTGATGT-3′). PCR conditions included initial denaturation (95° for 5 min), followed by 30 cycles of denaturation (95° for 30 sec), annealing (60° for 40 sec), and extension (72° for 2 min), with a final extension (72° for 10 min). The PCR products were cloned into pCR 2.1–TOPO (Invitrogen), and sequenced with Termination Reaction Ready Mix (Applied Biosystems) on an automated DNA Sequencer. This procedure yielded clones containing baboon apo[a] and plasminogen gene sequences. Three different apo[a] clones and two different plasminogen clones were sequenced on both strands using overlapping primers in each direction.

To determine genomic sequences of the transcript positive null allele and other apo[a] alleles, the baboon intron sequences were used to construct primers specific for either apo[a] or plasminogen genes (exons F, G, and H and intervening introns). The gene-specific primers were used for amplification of genomic DNA, and the PCR products were directly sequenced by cycle sequencing on an automated DNA sequencer (see Fig. 3 for PCR and sequencing primers). PCR conditions included initial denaturation (95° for 5 min), followed by 30 cycles of denaturation (95° for 30 sec), annealing (68° for 40 sec), and extension (75° for 2 min), with a final extension (72° for 10 min).

### PCR-based typing of the splice site mutation

After the splice site mutation was identified, a PCR-based assay was developed to type the mutation in genomic DNA from a total of 280 baboons. Genomic DNA samples were amplified with apo[a]-specific primers from the introns flanking exon G (primers given in Fig. 3). PCR conditions included initial denaturation (95° for 5 min), followed by 30 cycles of denaturation (95° for 30 sec) and annealing/extension (68° for 40 sec.) with no final extension step. The PCR products were digested with 1.0 U RsaI (Promega, Madison WI), and electrophoresed on agarose gels (3%). The wild type sequence contains the RsaI site

and yields 173 bp and 139 bp fragments, while the T mutation in the splice site disrupts the RsaI site yielding a full-length 312 bp fragment.

## RESULTS

### Baboon apo[a] and plasminogen cDNA sequences

We obtained the sequence of portions of baboon apo[a] and plasminogen cDNA using clones from hepatic cDNA libraries and from RT-PCR of hepatic mRNA. **Figure 1A** shows the sequence of the wild-type baboon apo[a] cDNA, including two Kringle IV domains (homologous to human Kringles type IV subtypes 9 and 10), the protease domain, and the 3′ UTR. Figure 1B shows the sequence of baboon plasminogen cDNA, including the Kringle V domain, protease domain, and 3′ UTR. We aligned the various domains of baboon cDNA sequences with those of other species, and found the highest homologies between baboon and rhesus cDNA sequences. **Table 1** shows pairwise comparisons among apo[a] and plasminogen domains from baboon, rhesus, and human cDNA sequences.

### Deletion of an exon in the protease domain in apo[a] mRNA of a transcript positive null allele

In addition to the wild-type sequence, we isolated and sequenced apo[a] cDNA from a baboon carrying an apo[a] null allele that produced hepatic transcripts (transcript positive null allele), and that was defective in secretion as shown by studies of cultured primary hepatocytes (12). This baboon was heterozygous for two different null alleles, one each from the sire and dam. Subsequent analyses (described below) showed that the dam transmitted the transcript positive null allele and the sire transmitted the transcript negative null allele which is not represented in cDNA analyses. The apo[a] genotypes for each family member of this baboon are shown in **Fig. 2A.**

Alignment of wild-type and null apo[a] sequences revealed the deletion of 141 base pairs (bp) in the null allele relative to the wild-type allele, exactly corresponding to exon G of the apo[a] gene (corresponding to exon 16 of the plasminogen gene (15)). This in-frame deletion results in the removal of 47 amino acids of the apo[a] protease domain. To confirm the presence of the deletion directly in hepatic apo[a] mRNA, we used RT-PCR of the protease domain in this family of baboons carrying null and wild-type alleles (Fig. 2B). In the baboon known to carry the transcript positive allele, we detected a shortened RT-PCR product corresponding to the 141 bp deletion detected in cDNA sequence comparisons. We found the dam was heterozygous for both the short and normal sized RT-PCR products, and other family members showed only normal sized products. We gel-isolated and sequenced the wild-type and shortened RT-PCR products, and confirmed that the size differences were due to the deletion of exon G. RT-PCR analysis of hepatic RNA samples from additional unrelated baboons (including 11 carrying transcript positive and 9 carrying transcript negative null alleles) failed to detect the short RT-PCR product.

### Identification of a splice site mutation in the intron downstream from the deleted exon

To investigate the source of the exon G deletion, we compared the exon and intron sequences in the protease domain of the wild-type versus transcript positive apo[a] null alleles. To obtain these sequences, we first determined genomic sequences in the protease domain of the baboon apo[a] and plasminogen genes. **Figure 3** shows the nucleotide sequences of exons F, G, and H (corresponding to plasminogen exons 15, 16, and 17 from ref. 15) plus their flanking introns for the apo[a] and plasminogen genes. We used intron sequences that differed between the two genes to construct apo[a]-specific primers for PCR and direct determination of genomic sequences from the baboon carrying the transcript positive null allele and a transcript negative null allele (Fig. 2A). We found that this baboon was heterozygous for an A→T mutation in the third nucleotide of the intron between exons G and H, disrupting a consensus splice site for transcript processing (Fig. 3). We sequenced this region of the apo[a] gene in the dam that was heterozygous for the shortened apo[a] mRNA (Fig. 2B), and found both the mutant T and wild-type A alleles. We also sequenced this region in 50 additional unrelated baboons (including 11 baboons carrying transcript positive null alleles and 9 carrying transcript negative null alleles), but did not find the A→T splice site mutation.

Nucleotide sequencing detected the splice site mutation only in baboons containing the transcript positive null allele and deleted apo[a] mRNA. In order to screen a larger number of baboons for the mutation, we developed a PCR-based assay using amplification of genomic DNA with apo[a]-specific intron primers, followed by digestion with RsaI which distinguishes the A→T mutation. The wild-type sequence contains an RsaI site that is abolished by the T mutation. We did not detect the A→T mutation in DNA samples from any of an additional 280 baboons that were typed using this assay.

## DISCUSSION

We used cDNA libraries and RT-PCR to clone and sequence a portion of the baboon apo[a] cDNA including two Kringle IV repeats, the protease domain, and the 3′ UTR (Fig. 1A). Sequence comparisons using the BLAST program (NCBI) showed that the Kringle IV repeats were most homologous to Kringles IV subtypes 9 and 10, which are the most 3′ kringles in human apo[a] cDNA (16). The baboon Kringle IV subtype 9 repeat contains a free cysteine that is required for covalent interaction with apoB-100 (17), and the baboon Kringle IV subtype 10 repeat contains a tryptophan (amino acid 171) important for lysine binding (18). Like the rhesus, baboon apo[a] cDNA does not contain a Kringle V domain that is found in human apo[a] cDNA and all other plasminogen cDNAs (13). Surprisingly, the baboon apo[a] protease domain, like plasminogen, contains arginine (position 224, Fig. 1A) at the cleavage site that converts inactive plasminogen

Panel A

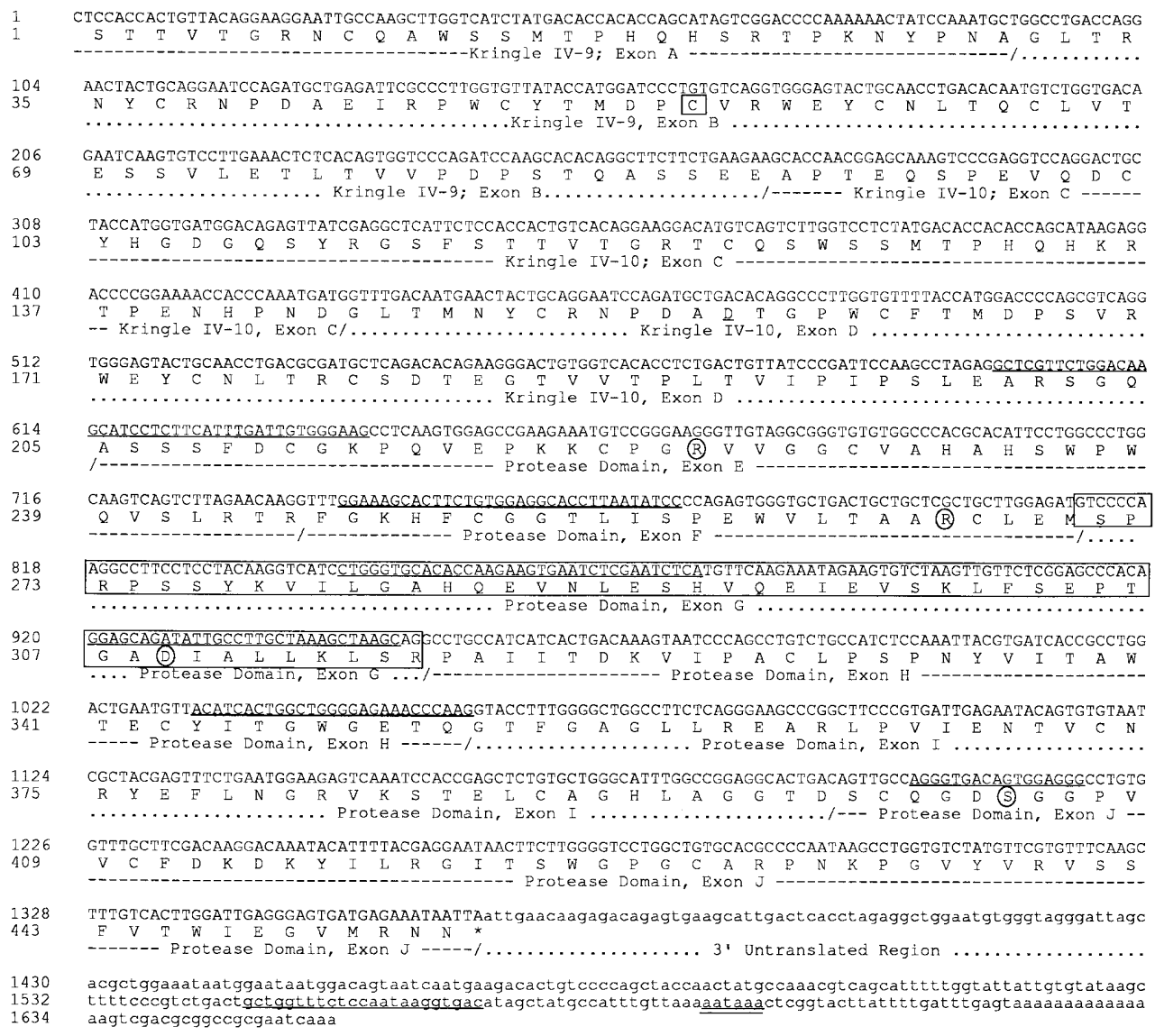

**Fig. 1.** Nucleotide and predicted amino acid sequences of portions of cDNA encoding baboon apolipoprotein [a] and plasminogen. Panel A shows the baboon apo[a] cDNA sequence. Nucleotide sequences (above) and amino acid sequences (below) are numbered to the left, and the exons are identified below the sequences (exon D corresponds to exon 14 in human plasminogen (15)). Exon structure within the protease domain is based on genomic sequences from the baboon apo[a] gene. The arginine at amino acid position 224 is circled, corresponding to the arginine in plasminogen which is cleaved to yield proteolytically active plasmin. Amino acids R (position 266), D (309), and S (404) corresponding to the catalytic triad of plasminogen (H,D,S) are also circled. Aspartic acid at position 157, which is underlined, is a potential lysine binding site. The free cysteine (position 54) in exon B thought to be necessary for apoB-100 binding is boxed. The 3′ untranslated region is shown in small case letters and the polyadenylation signal is double underlined. Exon G, which is deleted in the mRNA of the transcript positive null allele, is boxed. Primer sequences used for RT-PCR and sequencing are underlined. Genbank Accession Number: AF029691.

to active plasmin. However, one of the three amino acids (position 266, Fig. 1A) comprising the catalytic triad is an arginine instead of a histidine. Human apo[a] contains serine rather than arginine at the plasminogen cleavage site. However, a recent study by Gabel and Koschinsky (19) showed that reintroduction of arginine at the cleavage site does not restore proteolytic properties to recombinant human apo[a]. They concluded that proteolytic properties must require as yet unknown sequences in addition to the arginine cleavage site and catalytic triad. Although we have not yet tested baboon apo[a] for proteolytic activity, it seems likely that the protease domain that contains an altered catalytic triad will also prove to be inactive, despite the presence of the arginine cleavage site.

We cloned and sequenced a portion of the apo[a] cDNA for a transcript positive null allele that was previously shown to be defective in apo[a] secretion from pri-

Panel B

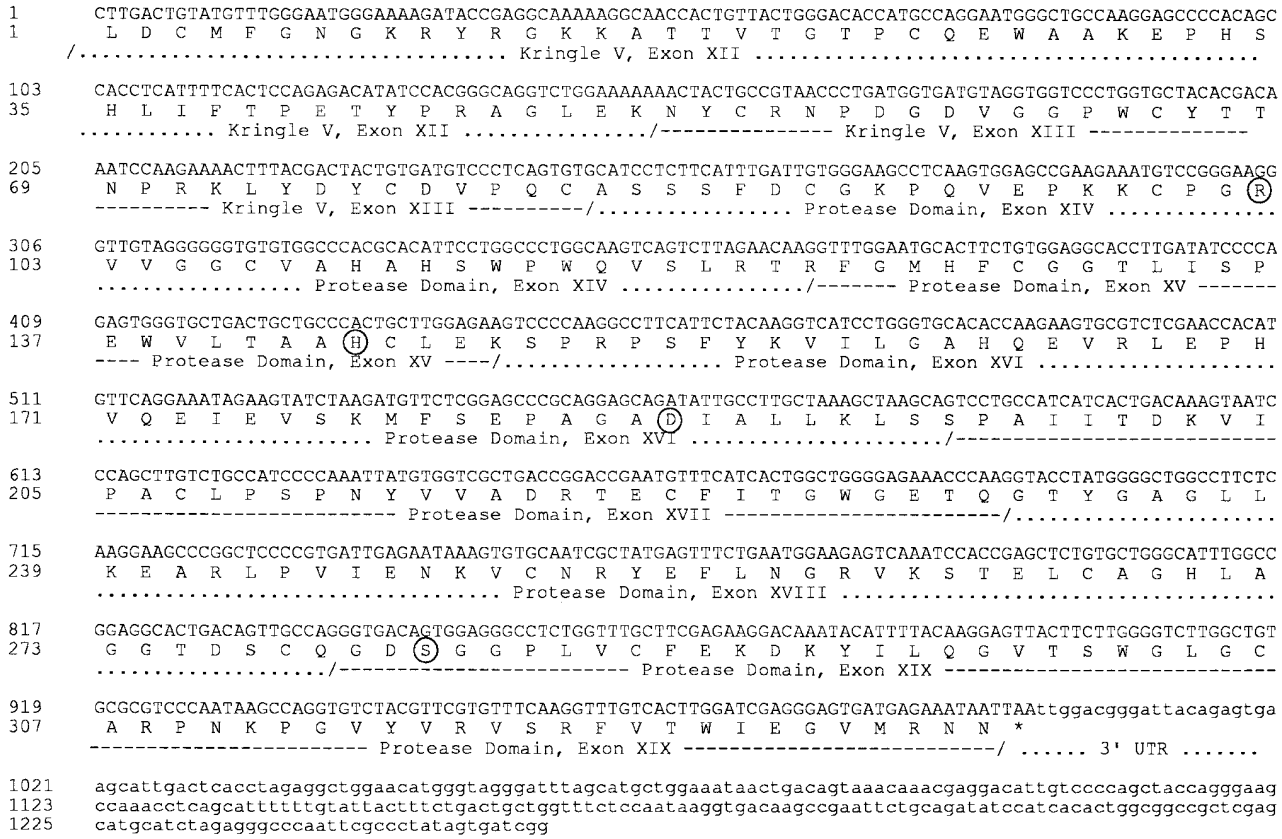


**Fig. 1.** (*continued*) Panel B shows the baboon plasminogen cDNA sequence. Nucleotide 1 corresponds to nucleotide 1491 of the rhesus plasminogen cDNA. The Kringle V, protease domain, and 3′ untranslated region are shown under the sequence. The arginine at amino acid position 102 is circled, corresponding to the arginine in human plasminogen which is cleaved to yield proteolytically active plasmin. Amino acids H (position 144), D (187), and S (282) are circled, corresponding to the plasminogen catalytic triad. The 3′ untranslated region is shown in small case letters. Genbank Accession Number: AF029692.

mary hepatocytes. We found a 141 bp deletion that removed a single exon in the protease domain (exon G), including one amino acid of the catalytic triad of the protease domain (Fig. 1A). The deletion was confirmed by RT-PCR of hepatic RNA, which yielded a shortened product relative to the wild-type apo[a] allele (Fig. 2B). We also found that the original baboon inherited the null allele from her dam who was heterozygous for the deletion (Fig. 2A). We determined genomic sequences in this region of the apo[a] gene for wild-type and transcript positive null apo[a] alleles, and found an A → T mutation in the third nucleotide position of the intron immediately downstream from the deleted exon G (Fig. 3). We found the T mutation only in the original baboon carrying the transcript positive null allele and her dam. RT-PCR screening of mRNA and genomic sequencing from an additional

TABLE 1. Homology comparisons among apo[a] and plasminogen sequences (% identities)

| | Kringle IV | | Kringle V | | Protease | | 3′ UTR |
|---|---|---|---|---|---|---|---|
| | DNA | Protein | DNA | Protein | DNA | Protein | DNA |
| Baboon apo[a] vs. rhesus apo[a][a] | 88 | 88 | N/A | N/A | 96 | 94 | 96 |
| Baboon apo[a] vs. human apo[a][b] | 86 | 87 | N/A | N/A | 92 | 92 | 85 |
| Baboon apo[a] vs. rhesus plasminogen[c] | 78 | 75 | N/A | N/A | 91 | 92 | 73 |
| Baboon apo[a] vs. human plasminogen[d] | 80 | 80 | N/A | N/A | 91 | 92 | 71 |
| Baboon apo[a] vs. baboon plasminogen | N/A | N/A | N/A | N/A | 94 | 94 | 77 |
| Baboon plasminogen vs. rhesus plasminogen | N/A | N/A | 97 | 96 | 98 | 97 | 96 |
| Baboon plasminogen vs. human plasminogen | N/A | N/A | 94 | 95 | 95 | 95 | 90 |

[a]Rhesus apo[a] sequences from ref. 13.
[b]Human apo[a] sequences from ref. 7.
[c]Rhesus plasminogen sequences from ref. 13.
[d]Human plasminogen sequences from ref. 34.

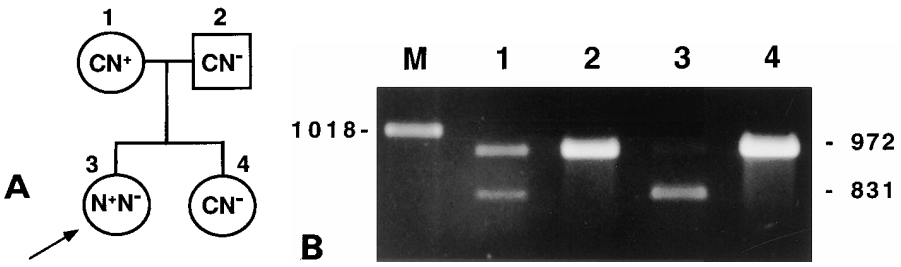


**Fig. 2.** Pedigree and RT-PCR products for the baboon family carrying the transcript positive null allele. Panel A shows the pedigree for the baboon (indicated by the arrow) carrying the transcript positive null allele that was previously shown to be defective with respect to apo[a] secretion from primary hepatocytes (12). The apo[a] genotypes are given adjacent to each symbol (C, single banded apo[a] phenotype type C; N+, transcript positive null allele; N−, transcript negative null allele). Each family member is numbered according to the lane number in Panel B. Panel B shows RT-PCR products of the apo[a] protease domain in the baboon family shown in panel A. Lane 1 shows RT-PCR products from dam, lane 2 shows the sire, lane 3 shows the baboon known to carry a transcript positive null allele, and lane 4 shows the sibling. The sizes of RT-PCR products are indicated to the right of the figure. Lane M shows the size standard (size shown on the left).

50 baboons (including 11 transcript positive and 9 transcript negative nulls), and genomic typing of 280 baboons did not detect the deletion or mutation, suggesting that this splice site mutation is relatively rare.

The A → T mutation in the third position of the intron occurs in the consensus site for recognition by components of the transcript splicing machinery, suggesting that the associated exon G deletion is due to exon skipping during mRNA processing. An explanation for this form of exon skipping is provided by the "exon definition model" of transcript processing in which the splicing machinery first scans and defines an exon based on strength and

```
          Exon F                A1F >                                                    intron f
Apo(a)  TTTGGAAAGCACTTCTGTGGAGGCACCTTAATATCCCCAGAGTGGGTGCTGACTGCTGCTCGCTGCTTGGAGAT|gtatgtttaactgacaggtggtctt
Plas    ---[---T-------------------G]----------------------------C-A-----------A ---------ggg----at--aca-g
                    P1F >

                                                                                                        A3R <
Apo(a)  atcgtcaagacttttttcctcttcctcttcctccctcttttgttccctctcgtcttcccctcttctcctccccacctccctttcctttctggaaggaaca
Plas    -ggt-ttgtct-aaa-a-t-t-----tcc-tcttt--cc-cc--t---t--tctc--tt-ctcc--t-cttttctgg-tg-ga-ac-agg-acca-gc--

Apo(a)  ctaggaaccaggaatgcacgcacagaa.....................Intron f.....................................................
Plas    gggca-t-agata-gg-ag.....................................................................................

                                    A2F >
Apo(a)  ttcttgccactcataagtcactgattctgaatggccaaggtctcaggaagattgtgccaaggtcatggcacagagggtacctgaaggggctggaccgta
Plas    ...----------g--------a----[-a-g-a--------g---g---ca-gat----t---c----------t----------]----------

                                    P2F >
                                  Exon G
Apo(a)  cttttctcttgacatcctcatcttttctag|GTTCTCAAGGCCTTCCTCCTATAAGGTCATCCTGGGTGCACACCAAGAAGTGAATCTCGAATCTCATGTT
Plas    t---g------------------------ --C-C----------A-T---C------------------------------CG-------C-A------

                                                                                    intron g
Apo(a)  CAAGAAATAGAAGTGTCTAGGTTGTTCTTCGGAGCCCACAGGAGCAGATATTGCCTTGCTAAAGCTAAGCAG|gtgctaacggtgctgtcagttaagaaaca
Plas    --G----------A----A-A------C-------G---------------------------------- -----cg-tcac----ggtc-tcaccct-

Apo(a)  gttatttgtaggagaggcaggtttgggacaggtgacaaggcacatagggcactcgctgtgctggcgactctggagggcagggtgtcggtgcagacaagga
Plas    tgctgg--a--atagttgctt-a--tctgg-t-ttatg--tcatggtct-tgcatggca-tg--gaggggct-tctatcaca--aaa-gct---ggggctt

          A1R <
Apo(a)  tggtctggcctggatgggaaagcatggggca.....................Intron g.....................................................
Plas    ---ggacag-atc-atcttc-]
          P1R <
                                    A3F >
Apo(a)  tgaaacgctgtgttctatgttgctctgtgtcattgaagcaaggcagtgccagctcagagggctctgagacctcaaggtaaggatgcctagttgtagatac
Plas    ----------------c--------------g-------------------------------c-g-g--------c-g-----g--g-----g---gt

                                                  Exon H
Apo(a)  tgcagctcccagctgggtagtcaaatataactgctaatacttcccttatag|GCCTGCCATCATCACTGACAAAGTAATCCCAGCCTGTCTGCCATCTCCA
Plas    -------g-----a-a-c-------c---------g--g---tt---tc-- T----------------------------------T----------C---

                                                P/A2R <
Apo(a)  AATTACGTGATCACCGCCTGGACTGAATGTTACATCACTGGCTGGGGAGAAACCCAAGgt
Plas    -----T---G--G-T-A-C----C-------T----------------------------
```

**Fig. 3.** Genomic DNA sequences of a portion of the baboon apo[a] gene aligned with the baboon plasminogen gene. The nucleotide sequences of exons F, G, and H (corresponding to exons 14, 15, and 16 of the plasminogen gene (15)) are shown in upper case letters, and the introns (f and g) are shown in lower case letters. Nucleotides in the baboon plasminogen sequence which differ from apo[a] are given below the apo[a] nucleotide sequence. The site of the A → T mutation in apo[a] intron g is circled, and the consensus splice site is double underlined. The primers for cloning and sequencing are boxed and labeled (A1F and A1R, primers for apo[a] fragment 1; A2F and P/A2R, apo[a] fragment 2; P1F and P1R, plasminogen fragment 1; P2F and P/A2R, plasminogen fragment 2; A3F and A3R, apo[a] sequencing primers).

proximity of processing sites, exon length, RNA secondary structure, exon enhancer elements, and intron sequences (reviewed in ref. 20). Thus, exons, rather that splice sites, provide the unit that is recognized for assembly of the splicesome (20, 21). After exon recognition, the splicing machinery cleaves introns and re-ligates the coding sequences, before transporting the mRNA to the cytoplasm for translation. Splice sites are determined by sequences present at the exon/intron junctions. Consensus sequences for eukaryotic splice sites include a donor splice site at the 5′ end of the intron ($^{C}/_{A}$ A G g u $^{a}/_{g}$ a g u) and a splice site at the 3′ end of the intron ($y_{10}$ n c a g $^{G}/_{A}$) (22, 23). The A → T mutation in this null apo[a] allele disrupts the 5′ donor splice site, replacing the $^{a}/_{g}$ with a uracil at the third position of the intron in the consensus sequence.

**Figure 4A** shows predicted normal splicing of exons F, G, and H in the apo[a] protease domain in which the exon ends are recognized by the splicing machinery, the introns between the exons are removed, and the ends of the exons are ligated together. Figure 4B shows predicted splicing of the null allele in which the splicing machinery fails to recognize both ends of exon G due to the A → T mutation at the 5′ donor splice site. Failure to recognize both ends of the exon and failure to recognize an alternative donor splice site may cause the splicing machinery to skip exon G as if it were an additional intron sequence between exons F and H (21, 24). Comparisons of known 5′ splice site mutations show that exon skipping is the most common outcome for this type of mutation (21). Other examples where intronic mutations at this position in the 5′ donor splice site lead to deletion of the preceding exon include LCAM deficiency (G → C) (25), eliptocytosis (G → T in the β-spectrin gene) (26), and other gene defects (24, 27). A splice site defect has also been detected in an apo[a]-related gene that causes exon skipping and a coding frame shift, resulting in a premature stop codon and truncation at the second kringle domain (28).

We have not characterized the entire null allele sequence and cannot formally exclude the possibility of some other sequence alteration that might affect apo[a] secretion. However, the exon G deletion was found only in the null allele that is known to be defective in secretion and in the same allele transmitted by the dam, but was not detected in any of the 330 additional baboons. Therefore, it is extremely unlikely that the deletion is a common polymorphism that by chance was associated with the null allele. Furthermore, the deletion removes a large portion of the protease domain which likely has adverse effects on normal protein folding and processing. The deletion of exon G removes 47 amino acids in the predicted β-chain of the protease domain (15), perhaps disrupting important interactions of cysteine residues for normal apo[a] folding. White et al. (12) did not detect any obvious folding defects in this null allele in cultured primary hepatocytes, but the resolution of those assays may not have been sufficient to detect such small differences in folding of the large apo[a] protein. However, such improper folding would be detected by cell processing machinery, resulting in retention in the endoplasmic reticulum and subsequent degradation (29, 30).

To our knowledge, this is the first report of a molecular basis for null alleles in apo[a]. Previous studies showed that two-thirds of null alleles in baboons do not produce hepatic apo[a] RNA (transcript negative nulls). Yet, in vitro transfection studies of 5′ flanking regions from a transcript negative null allele showed only small differences in promoter activities relative to wild-type alleles (31). Similarly, Bopp et al. (32) found that 5′ flanking regions of two human null alleles had active promoters in transfection experiments. It seems likely that mutations in other, as yet unknown, promoter elements must be re-

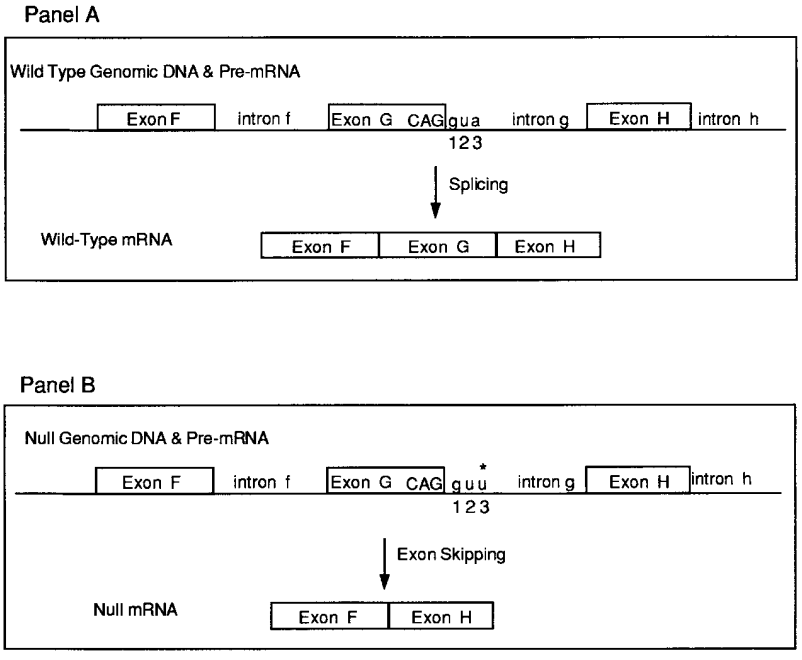


**Fig. 4.** Models for apo[a] transcript splicing in wild-type and the transcript positive null allele. Panel A shows a model for normal splicing of wild-type apo[a] mRNA where the ends of each exon are recognized by the splicesomal complex, the intron sequences are spliced out, and the exons are ligated together. Panel B shows a model for aberrant splicing due to the A → T mutation. The mutant uracil in the pre-mRNA in the consensus splice site of intron g (position 3) is shown by an asterisk. This disruption of the 5′ donor splice site may cause failure of the splicesome to recognize both ends of exon G, resulting in exon skipping and ligation of exons F and H.

sponsible for transcript negative nulls. Previous studies in baboons found that transcript positive null alleles produced different sized hepatic apo[a] transcripts, presumably resulting from different sized apo[a] alleles (10). Previous studies in humans found that null alleles were associated with different sizes of apo[a] alleles as measured by Southern blotting of genomic DNA (33). These results show that null alleles are not likely due to increased sizes of apo[a] alleles, which are often associated with lower Lp[a] levels. These results also indicate that null phenotypes occur on many different allelic backgrounds, perhaps due to many different mutations that interfere with apo[a] transcription, intracellular processing, or secretion. This hypothesis is further supported by the present study that shows a particular splice site mutation associated with defective secretion and a null phenotype that is rare in baboons, found only in a single baboon pedigree.

This research was supported by grant HL28972 from the National Institutes of Health. We thank Drs. Robert Lanford and Ann White who identified and provided a liver sample from the baboon carrying the transcript positive null allele. We also thank Dr. David Rainwater who determined the apo[a] phenotypes for the baboon population.

*Manuscript received 4 December 1997 and in revised form 5 March 1998.*

## REFERENCES

1. Walton, K. W., J. Hitchens, and H. N. Magnani. 1974. A study of methods of identification and estimation of Lp[a] lipoprotein and of its significance in health, hyperlipidaemia and atherosclerosis. *Atherosclerosis.* **20:** 323–346.
2. Hajjar, K. A., D. Gavish, J. L. Breslow, and R. L. Nachman. 1989. Lipoprotein[a] modulation of endothelial cell surface fibrinolysis and its potential role in atherosclerosis. *Nature.* **339:** 303–305.
3. Rhoads, G. G., G. Dahlen, K. Berg, N. E. Morton, and A. L. Dannenberg. 1986. Lp[a] lipoprotein as a risk factor for myocardial infarction. *J. Am. Med. Assoc.* **256:** 2540–2544.
4. Kostner, G. M., P. Avogaro, G. L. Cazzalato, E. Marth, G. Bittolo-Bon, and G. B. Quinci. 1981. Lipoprotein Lp[a] and the risk for myocardial infarction. *Atherosclerosis.* **38**: 51–61.
5. Boerwinkle, E., C. Leffert, J. Lin, C. Lackner, G. Chiesa, and H. Hobbs. 1992. Apolipoprotein[a] gene accounts for greater than 90% of the variation in plasma lipoprotein[a] concentrations. *J. Clin. Invest.* **90:** 52–60.
6. DeMeester, C., X. Bu, R. Gray, A. Lusis, and J. Rotter. 1995. Genetic variation in lipoprotein [a] levels in families enriched for coronary artery disease is determined almost entirely by the apolipoprotein[a] gene locus. *Am. J. Hum. Genet.* **56:** 287–93.
7. McLean, J. W., J. E. Tomlinson, W-J. Kuang, D. L. Eaton, E. Y. Chen, G. M. Fless, A. M. Scanu, and R. M. Lawn. 1987. cDNA sequence of human apolipoprotein[a] is homologous to plasminogen. *Nature.* **330:** 132–137.
8. Lackner, C., E. Boerwinkle, C. C. Leffert, T. Rahmig, and H. H. Hobbs. 1991. Molecular basis of apolipoprotein[a] isoform size heterogeneity as revealed by pulsed-field gel electrophoresis. *J. Clin. Invest.* **87:** 2153–2161.
9. Williams-Blangero, S., and D. L. Rainwater. 1991. Variation in Lp[a] levels and apo[a] isoform frequencies in five baboon species. *Hum. Biol.* **63:** 65–76.
10. Hixson, J. E., M. L. Britten, G. S. Manis, and D. L. Rainwater. 1989. Apolipoprotein [a] (Apo[a]) glycoprotein isoforms result from size differences in apo[a] mRNA in baboons. *J. Biol. Chem.* **264:** 6013–6016.
11. White, A. L., D. L. Rainwater, J. E. Hixson, L. E. Estlack, and R. E. Lanford. 1994. Intracellular processing of apo[a] in primary baboon hepatocytes. *Chem. Phys. Lipid.* **67/68**: 123–133.
12. White, A., B. Guerra, and R. Lanford. 1997. Influence of allelic variation on apolipoprotein[a] folding in the endoplasmic reticulum. *J. Biol. Chem.* **272:** 5048–5055.
13. Tomlinson, J., J. McLean, and R. Lawn. 1989. Rhesus monkey apolopoprotein[a], sequence, evolution and sites of synthesis. *J. Biol. Chem.* **264:** 5957–5965.
14. Ichinose, A. 1992. Multiple members of the plasminogen-apolipoprotein[a] gene family associated with thrombosis. *Biochemistry.* **31:** 3113–3125.
15. Petersen, T. E., M. R. Martzen, A. Ichinose, and E. W. Davie. 1990. Characterization of the gene for human plasminogen, a key proenzyme in the fibrinolytic system. *J. Biol. Chem.* **265:** 6104–6111.
16. Koschinsky, M., J. Tomlinson, T. Zioncheck, K. Schwartz, D. Eaton, and R. Lawn. 1991. Apolipoprotein[a]: expression and characterization of a recombinant form of the protein in mammalian cells. *Biochemistry.* **30:** 5044–5051.
17. Koschinsky, M., G. Cote, B. Gabel, and Y. van der Hoek. 1993. Identification of the cysteine residue in apolipoprotein[a] that mediates extracellular coupling with apolipoprotein B-100. *J. Biol. Chem.* **268:** 19819–19825.
18. LoGrasso, P. V., S. Cornell-Kennon, and B. R. Boettcher. 1994. Cloning, expression, and characterization of human apolipoprotein[a] Kringle IV37. *J. Biol. Chem.* **269:** 21820–21827.
19. Gabel, B., and M. Koschinsky. 1995. Analysis of the proteolytic activity of a recombinant form of apolipoprotein[a]. *Biochemistry.* **34:** 15777–15784.
20. Zandberg, H., T. Moen, and P. Baas. 1995. Cooperation of 5′ and 3′ processing sites as well as intron exon sequences in calcitonin exon recognition. *Nucleic Acids Res.* **23:** 248–255.
21. Talerico, M., and S. Berget. 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. *Mol. Cell Biol.* **10:** 6299–6305.
22. Shapiro, M., and P. Senapathy. 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15:** 7155–7174.
23. Mount, S. 1982. A catalogue of splice junction sequences. *Nucleic Acids. Res.* **10:** 459–472.
24. Carstens, R., W. Fenton, and L. Rosenberg. 1991. Identification of RNA splicing errors resulting in human ornithine transcarbamylase deficiency. *Am. J. Hum. Genet.* **48:** 1105–1114.
25. Kishimoto, T., K. O'Connor, and T. Springer. 1989. Leukocyte adhesion deficiency: aberrant splicing of a conserved integrin sequence causes a moderate deficiency phenotype. *J. Biol. Chem.* **264:** 3588–3595.
26. Garbarz, M., W. Tse, P. Gallagher, C. Picat, M.-C. Lecomte, F. Galibert, D. Dhermy, and B. Forget. 1991. Spectrin Rouen ($\beta^{220\text{-}218}$), a novel shortend β-chain variant in a kindred with hereditary elliptocytosis. *J. Clin. Invest.* **88:** 76–81.
27. Krawczak, M., J. Reiss, and D. Cooper. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* **90:** 41–54.
28. Byrne, C., K. Schwartz, K. Meer, J-F. Cheng, and R. Lawn. 1994. The human apolipoprotein[a]/plasminogen gene cluster contains a novel homologue transcribed in liver. *Arterioscler. Thromb.* **14:** 534–541.
29. Bergeron, J., M. Brenner, D. Thomas, and D. Williams. 1994. Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends Biochem. Sci.*. **19**: 124–128.
30. Hartl, F. 1996. Molecular chaperones in cellular protein folding. *Nature.* **381:** 571–579.
31. Hixson, J. E., C. Jett, and S. Birnbaum. 1996. Identification of promoter sequences in the 5′ untranslated region of the baboon apolipoprotein [a] gene. *J. Lipid Res.* **37:** 2324–2331.
32. Bopp, S., S. Kochl, F. Acquati, P. Magnaghi, A. Petho-Schramm, H-G. Kraft, G. Utermann, H-J. Muller, and R. Taramelli. 1995. Ten allelic apolipoprotein[a] 5′ flanking fragments exhibit comparable promoter activities in HepG2 Cells. *J. Lipid Res.* **36:** 1721–1728.
33. Gaw, A., E. Boerwinkle, J. Cohen, and H. Hobbs. 1994. Comparative analysis of the apo[a] gene, apo[a] glycoprotein, and plasma concentrations of Lp[a] in three ethnic groups. *J. Clin. Invest.* **93:** 2526–2534.
34. Forsgren, M., B. Raden, M. Israelsson, K. Larsson, and L-O. Heden. 1987. Molecular cloning and characterization of a full-length cDNA clone for human plasminogen. *FEBS Lett.* **213:** 254–260.