Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 1369–1373, February 1993
Biochemistry

5′ control regions of the apolipoprotein(a) gene and members of the related plasminogen gene family

[apolipoprotein/atherosclerosis/lipoprotein(a)]

DAVID P. WADE*, JOHN G. CLARKE*, GISELA E. LINDAHL*, ALEXANDER C. LIU*, BERNICE R. ZYSOW*, KELLI MEER*, KAREN SCHWARTZ†, AND RICHARD M. LAWN*

*Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305-5246; and †Arris Pharmaceutical Corp., 385 Oyster Point Boulevard, South San Francisco, CA 94080

Communicated by Charles G. Sibley, November 30, 1992

ABSTRACT **Elevated blood levels of apolipoprotein(a), the component of lipoprotein(a) that distinguishes it from low density lipoprotein, are a major risk factor for atherosclerosis. The apolipoprotein(a) gene is highly similar to the plasminogen gene and to at least four other genes or pseudogenes. The 5′ untranslated and flanking sequences of these six genes contain extensive regions of near identity and share sequence elements involved in the initiation of transcription and translation. About 1000 base pairs of flanking DNA of each gene are sufficient to promote transcription in cultured hepatocytes. The apolipoprotein(a) gene promoter contains functional interleukin 6-responsive elements, consistent with the reported acute-phase response of apolipoprotein(a). Flanking genomic fragments of the apoliprotein(a) gene from two individuals with vastly different plasma apolipoprotein(a) concentrations have sequence differences that are reflected in differences in the rate of *in vitro* transcription.**

Over the past decade, high plasma levels of lipoprotein(a) [Lp(a)] have been established as one of the major risk factors for atherosclerosis and its major manifestations: myocardial infarction, restenosis following angioplasty and bypass surgery, and stroke (for reviews, see refs. 1–3). Two independent studies conclude that 27% of premature myocardial infarctions in males may be due to elevated Lp(a) (4, 5). Unlike other lipoproteins, individual variation in Lp(a) concentration is almost entirely due to inheritance (6). Lp(a) resembles low density lipoprotein (LDL) in its content of cholesterol, phospholipids, and apolipoprotein B-100. It is distinguished from LDL by the large glycoprotein apolipoprotein(a) [apo(a)] that is attached to the particle by disulfide linkage to apolipoprotein B-100. Cloning and sequencing of the cDNA of human apo(a) revealed an unexpected homology to plasminogen (7). The 5′ regions of the two cDNAs have nearly 100% identity, whereas other domains of the plasminogen and apo(a) have between 75% and 94% local sequence identity. An unusual feature of the apo(a) cDNA is the large number of times that one domain is repeated. This domain is homologous to kringle four of plasminogen and is variously repeated from about 15 to 40 times in individual apo(a) genes, resulting in a large number of molecular weight isoforms of the protein (7–9). It is likely that the apo(a) gene arose from a duplication of the plasminogen gene, perhaps recently during mammalian evolution, with regions at the 5′ end undergoing subsequent gene conversion.

Plasma Lp(a) concentration varies widely in the human population, ranging from <0.1 to >200 mg/dl. Since blood levels of Lp(a) are a major contributor to morbidity, it is important to understand its regulation. Metabolic and genetic evidence shows that Lp(a) synthesis, rather than catabolism, is the key determinant for plasma concentration (10, 11) and that it is the apo(a) component of the lipoprotein particle that is the most important determinant of its final concentration (1, 6, 8). Studies in human populations show a general inverse relationship between the size of the apo(a) protein and plasma concentration (cf. refs. 1, 12, and 13). Protein translation or particle assembly might in some way account for this relationship. However, even within a given apo(a) isoform size class, a great deal of variation in plasma concentration exists, emphasizing the importance of distinct regulatory elements in the gene. This is supported by studies in humans and cynomolgous monkeys which show that hepatic apo(a) mRNA levels are related to Lp(a) concentration (14–16). To begin to understand the control of apo(a) gene expression, we present here the sequence of the flanking regions of the apo(a) gene and closely related homologues and demonstrate the activity of these regions in transcription assays.‡

MATERIALS AND METHODS

Cloning of 5′ Flanking Region of the Apo(a) Gene and Related Members of the Plasminogen Gene Family. Electrophoresis, blot hybridization, and other standard molecular biology procedures essentially followed the protocols of Sambrook *et al.* (17). Genomic libraries in bacteriophage λ (λ-ZAP II and Gigapack Gold II; Stratagene) were constructed from human genomic DNA digested to completion with *Eco*RI and screened with the probe pcr A (see below). In some cases, the *Eco*RI-digested DNA was size-selected by fractionation by 1% agarose gel electrophoresis. Positive clones were isolated as pBluescript plasmids by the *in vivo* excision procedures detailed by the manufacturer and sequenced to completion by the Sanger dideoxy method using a Sequenase kit (United States Biochemical) and oligonucleotide primers derived from either vector sequences or from previous sequence analysis of the cloned inserts. In some cases, nested deletion sets were produced by the exonuclease III mung bean system (Stratagene).

Oligonucleotide Primer Sequences. pcr 6, GCCTGTTGGAAAGCTTG, and pcr 7, AGTAGAAGAACCACTTC, were used to generate the genomic DNA 483-bp probe designated pcr A. ACCACATGGCTTTGC, the reverse complement of sequence in the first kringle region of apo(a) mRNA, was used to determine the mRNA start site by reverse transcriptase extension. Human liver poly(A)$^+$ RNA was isolated by the guanidium isothiocyanate procedure (18).

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.

Abbreviations: apo(a), apolipoprotein(a); LDL, low density lipoprotein; Lp(a), lipoprotein(a).

‡The sequences reported in this paper have been deposited in the GenBank data base [accession nos. L07899 for the human apo(a) gene exon 1, L07900 for apo(a)rg-B, and L07901 for apo(a)rg-C].

Luciferase Reporter Gene Constructs. To assay promoter activity, we synthesized a series of reporter gene constructs (pLuc A–D), comprising the 5′ flanking regions of the apo(a) gene and its homologues inserted into the polylinker region of pGeneLight-2 (Promega), a promoterless vector containing the firefly luciferase gene. The genomic inserts were isolated from modified pBluescript plasmids of the original clones from which genomic fragments 3′ to the desired flanking regions had been removed by restriction enzyme digestion. The flanking region inserts were isolated from these modified plasmids by digestion with restriction enzymes appropriate for directional ligation into the pGeneLight-2 polylinker. The inserts contained 1.4 kb of DNA upstream of a *Bal* I site located 12 nucleotides before the translation initiation ATG for the A1, A2, and B fragments, 1.07 kb upstream of this site in the D fragments, and 3.0 kb upstream of a *Hin*cII site at −54 for the C fragment.

Transfection and Assay of Reporter Gene Constructs. Plasmids for transfection were purified by two cesium chloride density gradient ultracentrifugations or by Qiagen (Studio City, CA) columns. For transient transfections with luciferase construct plasmids, HepG2 cells were seeded at 0.75 × 10^6 cells per 6-cm dish and maintained in Dulbecco's modified Eagle's medium (DMEM) (GIBCO/BRL) supplemented with 10% fetal calf serum. Transfections were performed when the cells attained 40–60% confluence. For each dish, 10 μg of test plasmid DNA and 5 μg of pSV-β-galactosidase control plasmid (Promega) were mixed with 1.5 ml of Opti-Mem medium (GIBCO/BRL) in a polystyrene tube, and 1.5 ml of Opti-Mem containing 30 μg of Lipofectin (GIBCO/BRL) was added. Complexes were allowed to form for 30 min at room temperature and the lipofectin/DNA mix was added to the dishes after washing the monolayers three times in Pucks saline A (GIBCO/BRL). After 16 hr the medium was replaced by 5 ml of DMEM containing 10% fetal calf serum, and incubation was continued for 48 hr before assay of cellular luciferase and β-galactosidase activity.

Monolayers were washed five times with ice-cold phosphate-buffered saline (PBS), scraped into 1 ml of PBS, and pelleted by centrifugation at 4°C for 10 sec in a Microfuge. The pellet was resuspended in 100 μl of 0.1 M potassium phosphate buffer, pH 7.8/1 mM dithiothreitol and lysed by three freeze-thaw cycles with vigorous Vortex mixing between each cycle. Lysates were stored in aliquots at −70°C if not assayed immediately. Luciferase activity was measured by mixing 100 μl of luciferase substrate solution (Promega) with 20 μl of cell lysate and a 0- to 10-sec integral was read in a Monolight 2001 luminometer (Analytical Luminescence Laboratories, San Diego). β-Galactosidase activity was measured after dilution of the lysates 1:100 in potassium phosphate buffer. Diluted lysate (20 μl) was incubated for 1 hr at room temperature with 200 μl of a 10 μg/ml AMPGD chemiluminescent β-galactosidase substrate (Tropix, Bedford, MA) in 0.1 M potassium phosphate/1 mM $MgCl_2$, pH 7.8. Light emission was initiated by addition of 300 μl of 1 M NaOH containing 10% volume-to-volume Emerald enhancer (Tropix), and a 0- to 10-sec integral was read immediately in a luminometer. Luciferase activities were normalized to β-galactosidase activity for each dish. Within each experiment luciferase activity was determined in duplicate dishes, and each plasmid was tested in three to five separate experiments; therefore results are expressed as the mean of 6–10 values.

RESULTS

Identification of Six Closely Related Genes or Gene Fragments. cDNA sequence analysis had shown that the 5′ ends of human apo(a) and plasminogen mRNA are nearly identical. Initial hybridization of human genomic DNA Southern blots with fragments from the 5′ region of the human apo(a) cDNA indicated the existence of several distinct genomic homologues of this region. A 483-bp probe (designated pcr A) covering this region was synthesized by PCR from genomic DNA, based upon the plasminogen gene sequence of Peterson *et al.* (19). When hybridized to an *Eco*RI digest of human DNA at high stringency, the pcr A probe detects four fragments designated A (4.8 kb), B (3.6 kb), C (3.1 kb), and D (1.3 kb) (Fig. 1 *Inset*). The intensity of the D fragment indicated that it might contain two or more copies. Several clones corresponding to each of these hybridizing genomic fragments were isolated from *Eco*RI complete-digest λ librar-

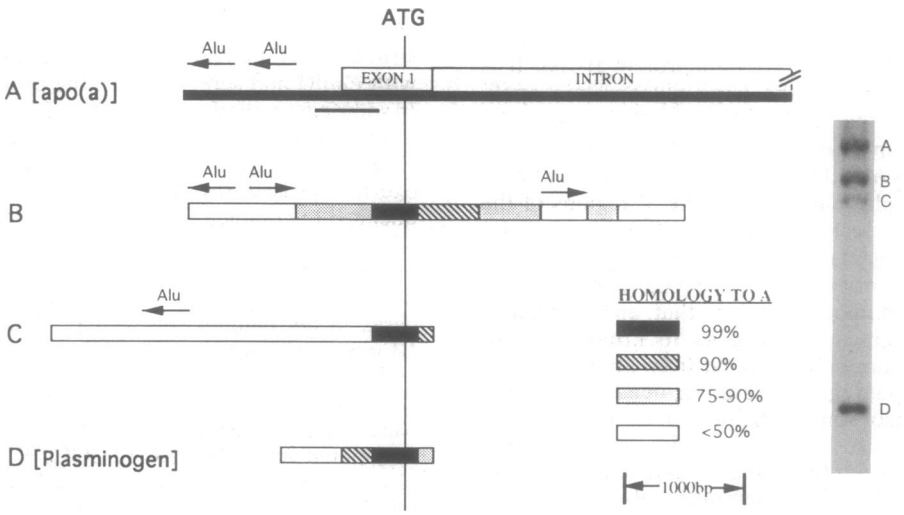

FIG. 1. Homologous regions from the 5′ ends of members of the apo(a)/plasminogen gene family. On the right is seen a blot hybridization of human genomic DNA digested with *Eco*RI, electrophoresed, transferred, and probed with pcr A, a 483-bp fragment from the 5′ end of the plasminogen gene. The hybridizing fragments are 4.8, 3.6, 3.1, and 1.3 kb in length. Local regions of DNA sequence identity to the 4.8-kb A fragment are shown. The fragments are shown aligned to the ATG translation initiation codon. Exon one begins at the transcription start site and ends at a splice donor sequence shared by all four fragments. The location of repetitious "*Alu*" sequences and their orientation are shown (arrows). The D fragment size class consists of three or more nonallelic species containing about 96% sequence identity throughout their length, one of which corresponds to the plasminogen gene. [The complete sequences of the 4.8-kb A fragment, designated human apo(a) gene exon 1; the 3.6-kb B fragment, designated apo(a)rg-B; and the 3.1-kb C fragment, designated apo(a)rg-C, have been deposited with GenBank.]

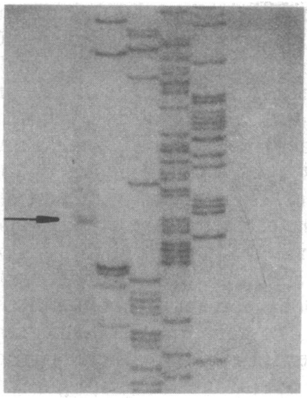

FIG. 2. Mapping of the transcription start site (arrow) of apo(a). An antisense primer complimentary to nucleotides 60–72 downstream of the ATG of human apo(a) cDNA (7) was annealed to 20 μg of human liver poly(A)$^+$ RNA and extended with reverse transcriptase as described (23). A DNA sequence ladder is included as a size marker.

ies and sequenced to completion. Sequence analysis confirmed the existence of more than four extremely homologous fragments of the genome, since three variants of the 1.3-kb fragment were isolated from one individual (Fig. 1). Comparing the sequence of the cloned genomic fragments to the published cDNA sequence of human apo(a) and plasminogen (7, 20, 21), we found that all four fragments contain a region of >95% sequence identity that encompasses the translation initiation codon and first exon of these genes and contains the identical sequence CAG/GTA at the first splice donor site. [The first exon of the human plasminogen and apo(a) genes ends at an identical location 49 nucleotides following the ATG and codes for the first 16 amino acids of the signal peptide sequence.] The 4798 bp of our A fragment sequence contain an exact match with the 400-bp sequence of apo(a) I given by Ichinose (22), and but one mismatch (at −21 before the ATG) with the overlapping region of apo(a) mRNA reported by McLean *et al.* (7) and extended by sequencing a pcr product of mRNA to nucleotide −141. The assignment of the A fragment to the *bona fide* apo(a) gene is consistent with the recovery of overlapping λ genomic clones that contain the A fragment sequence with clones encoding kringle regions of apo(a) (R.M.L., unpublished data). The transcription start site of apo(a) mRNA was determined by primer extension with reverse transcriptase (Fig. 2). The major start site appears to occur at 141 bases before the translation start ATG. This is 20 bases further downstream than the transcription start site reported for human plasminogen mRNA by Malgaretti *et al.* (24). [pcr analysis suggests that both plasminogen and apo(a) mRNAs have minor start sites further upstream.] The sequence of the 5′ region of the human apo(a) gene is given in Fig. 3. [We designate the B and C fragments as belonging to apo(a)-related genes, apo(a)rg-B and apo(a)rg-C.]

Promoter Activity of Genomic Fragments. To begin analysis of elements that control transcription of the apo(a) gene and to determine which of the other promoter fragments are functional, we inserted the cloned genomic fragments into luciferase reporter plasmids. Genomic DNA from the 5′ *Eco*RI site of each fragment to a point in the 5′ untranslated

```
                                                           GAATTCATTTGCGGAAAGATTGATACTATGCTTTTATTTTAT
                                  v
-1400  TTTATTTTATTTTATTTTATTTTATTTTATTGAGACTCTCACCCCGGTTGAAGTGCACTGACGTGATTTTGGCTCACTGCAACTTCCACCTCCTGGGTTC
                                 IL-6
-1300  AAGTGAATACTCCAGCCTCCCTAGTAGCTGGGATTACAGGTGCCCACCACCACGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCACAT
                                                                  IL-6          LF-A1
-1200  TGGCCTGGCTGGTCTCAAACTCCTGACCTTGTGATCCACCTGTCTTGGCCTCCCAAAGTGCTGGGATTACAGAGTTGAGCCACCGCACTCGACCCTATGT
-1100  TTTATTTTTAAAAATATTTATTTATTTATTTAAGCCACAACTACTAGAATAGGAAGGATTGATATTTTATTAATTTTATTTGGTATTTATTATTTTTTTT
                                                                                                 * CEBP
-1000  TCTTTCCTGAGACATTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGCACATTCTTGGCTCACTGCAACCTCCATCTCCTGTGTTCGAGCAATTCTAGT
                          IL-6
 -900  GCCTCAGCCTACTTAGTAGCTGGGATGACTGGCATGTGCCTCCACACCCAGCTAATTTTTGTATTTTTTGTAGAGACAGGGTTTTGGCATGTTGCCAGGC
                                                           IL-6
 -800  TTGTCTCAAACTCCTGGCCTCAGGTGATCCATCTGCCGTGGCCTCCCAAAATGCTGGGATTATAGGCATGAGCCACCACCCCCTCCTGGAAGGATTGATA
 -700  TCTTATAACATAATTTATAATTACAGAAAACATGTGAGTTCACTAGGAATAAATAAATTTTGAAGATAATAAAAGATTTTCACTTATGTTGTCATTTCGG
 -600  CACAGTTTGGTATAGGATGTGGAGATGTTAACATTTATACCTAGCTTGCTCGTAAACTAAGACCTGAAAGGGTTGTGTCTATCAGCTGCACCCCTGGGTA
 -500  GCGACACAACCTCGGGAAGGCCTCAGCCCCCTCCTCGTACAGCACTGCCTGTTGGAAAGCTTGAGGGAGGCTATGGATGTGCAGCACTTGGCAGAGGGTC
 -400  TGGTCATGGAAGTTACCAGCAAATATGAGCTACTTTTATGATTTTATTTTATCCAAAAGAAAGAGAATGAAAGAAGAGGGGAGGAAACAAGACTAATCAG
                           IL-6                    CEBP                         IL-6RC
 -300  GAAAGATGAAGGTCTAGGGGTGAGGGAAGGAGTAAGGAGACATAAAGGCAATGTGGAGCAGCTGAGGGGGGAAATGGCTTTCACCACTTCCCAGCATCTA
            CEBP-RC                                         LF-A1                                HNF-1α
 -200  TTGACATTGCACTCTCAAATATTTTATAAGACTCTATATTCAAGGTAATGTTTGAACCCTGCTGAGCCAGTGGCATGGGTCTCTGAGAGAATCATTAACT
                                                             IL-6
 -100  TAATTTGACTATCTGGTTTGTGGGTGCGTTTACTCTCATGTAAGTCAACAACGTCCTGGGATTGGACACACTTTCTGGGCACTGCTGGCCAGTCCCAAA
       Met
       ATGGAACATAAGGAAGTGGTTCTTCTACTTCTTTTATTTCTGAAATCAGGTAAGACATAGTTTTTTAAATTATAAGAATTATTTTTTCTCCCACAATGT

  101  AGTAAAAATACATATGCCATGGCTTTATGTGCAATTCATTTAATTTTTGATTCATGAAATTCCCAGTTCAAAATCTTGTATATGATTGAAAAATTCTTCC
  201  AAAAATTAGTTTAATTTCCCCGTGAAGACTGTCACGGTGCTGGAATGAATGGGCAGAAAAAATAATGGTTGATTTTTCTAATCTAAAAGAGTGTGCCTAC
  301  ATGATGGCCAGTCTGGCTGAAAAATAAATAGCCATTGTAGCTAACTATGCAAAGGATGGCTAAGCTCTTCGCTTGGTTCTCAGTTTCATTAATTTATATC
  401  ATCTCTGTTCAGGTGCCATGCTCCCCTCACTAGCAAGTTGAAACAATGAAATAACTCTTTGAATATGTTTGGTTCCTTGACCTGTTCATGGAGTGGGACT
  501  CAGCATTTCTCTCTTTGTTATGGCCTGAGTAAGGCTTTCCATCGGTATACATTTGCTTCTTATCCCTGGAGAAATTATACACATCCATTTGCCAGATGAT
  601  ATACACATATAATGATTCAACAAATACTCAGGGTATTTGTTGAGTGGGTTAGGTCCCCACATTTTTATACATACATACACACATACACACCGTGTGTGAT
  701  TGTGAATGTAAGTGTGTGTCCTTTACAAATACTAGCTTATTTAGCTCATGGTATAGGTAGGGTAGCATAGTCATCCCCATTTTATAAACAAAGAAATCTA
  801  GACTTAGGAAAATCATGTTATTTGTCTCGTGACCAAATTCCCAAATCAAGGAAATAAAGAAACCTGGATTTAAGCCAGATTTCCAAGAAAAAATCTAGGG
  901  CTCTTCTCACTTTTTCATCTTTGTTCCAACATTTGAAAAAATAAATCTAAACACATTCCAATGTAACTGAAGAGCAGGTTAATTGTTTGCCACTTGCAGA
 1001  ATCCAATTAAGAAGAGAGAAGTCTGGTATAAAGAAAGTGATTTGCTTCCAAAGCTAGCTTAGGGGAAGAAATGCAGCAGTCCTGCCGTACCACTTCACTT
```

FIG. 3. The 5′ region of the human apo(a) gene. Partial sequence of the *Eco*RI "A" fragment is presented; the complete 4798-bp sequence is available from GenBank. The first nucleotide of the translation initiation codon is numbered one. Sequences resembling CAAT, TATAA boxes, and other consensus sequences for transcription regulatory elements are boxed, and exon one is underlined. The A2 allele isolated from the individual with the low plasma and reporter gene expression of apo(a) (see Fig. 4 and text) has two sequence differences with the above: an insert of TTTTA indicated by v and a G → A substitution indicated by *. IL-6, interleukin 6.

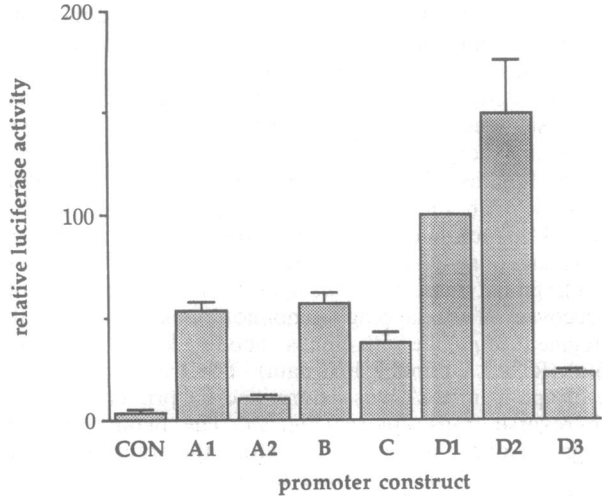

FIG. 4. Transcription activity of the homologous genomic fragments. The 5′ flanking regions of the genomic fragments were ligated just upstream to the coding region of the luciferase gene, transfected into HepG2 cells, and assayed for luciferase activity. Values are the result of three to five independent experiments and are presented relative to the activity of the D1 fragment. CON is the luciferase plasmid pGeneLight-2 without insert; A1 is the apo(a) genomic fragment derived from an individual with 30 mg of plasma apo(a) per dl; A2 is the corresponding fragment from an individual with <1 mg of plasma apo(a) per dl; D1, D2, and D3 derive from three distinct plasminogen-like 1.3-kb *Eco*RI genomic fragments isolated from a single individual.

region was inserted into the GeneLight-2 plasmid (Promega) upstream of the firefly luciferase gene (see *Materials and Methods*). The recombinant plasmids contained 1.4 kb of A fragment 5′ untranslated sequence, 1.2 kb of B fragment sequence, 3 kb of C fragment sequence, and 1.1 kb of D fragment sequence. All of the fragments proved competent to drive transcription of the reporter gene following transfection into HepG2 hepatocarcinoma cells (Fig. 4), suggesting that more than just plasminogen and apo(a) genes might be transcribed from this gene family (see *Discussion*). Apo(a) gene A fragments were cloned and tested from two individuals. A1 derived from an individual with relatively high plasma Lp(a) concentration (30 mg/ml), whereas A2 derived from an individual with Lp(a) below the level of detection by ELISA and by Northern blot hybridization. The only DNA differences in the two A fragments are a single base substitution 914 nucleotides before the ATG and nine, rather than eight, copies of the repeated sequence TTTTA at position −1373 (see Fig. 3; these two locations are not homologous to the D fragment sequence). When normalized to a value of 100% for the D1 fragment, the promoter activity of A1 was 53%, whereas A2 was only 11%, which is consistent with the lower plasma apo(a) concentration in the A2 individual. In addition, fragment B had 57% relative promoter activity, fragment C had 38%, D2 had 149%, and D3 had 23%.

DISCUSSION

Studies of the apo(a) gene have been complicated by the existence of multiple genomic elements with great sequence similarity. This includes the highly repeated multiple kringle domains of apo(a), with virtually 100% identity, and the six or more genes and/or pseudogenes with >95% sequence identity between portions of themselves. Apo(a) belongs to an extended gene superfamily that contain domains that resemble the serine protease trypsin. By addition, loss, and exchange of domains such as "kringles," epidermal growth factor precursor domains, fibronectin "fingers," and calcium binding domains, as well as by base substitution, superfamily members have evolved to produce genes encoding blood-clotting factors, plasminogen activitors, structural proteins, and growth factors (25, 26). One branch of the family tree that contains only kringle and protease-like domains includes plasminogen, apo(a), and hepatocyte growth factor. Hepatocyte growth factor contains 38% amino acid identity to plasminogen, whereas plasminogen and apo(a) are much closer homologues. Comparison of human plasminogen and apo(a) cDNA suggested that the apo(a) gene arose from a duplicated plasminogen gene through deletion of the exons encoding the preactivation peptide and kringles one through three and the multiplication of exons corresponding to its kringle four (7). The degree of sequence similarity and the virtual confinement of apo(a) to primates indicates the recent emergence of the apo(a) gene. One intriguing exception is the finding of an Lp(a)-like lipoprotein in the insectivore, the hedgehog (27). This may be an example of convergent evolution.

Hybridization to human chromosomes *in situ*, and to blots of human–rodent somatic cell hybrids, demonstrated that the highly related homologues of apo(a) and plasminogen genes are closely linked in the region 6q27 (28–30). This proximity is consistent with a history of duplication and gene conversion events. All of the homologous but distinct genomic fragments sequenced in this study contain elements that resemble 5′ flanking sequences, the first exon, and portions of the first intron of plasminogen and apo(a) genes. Peterson *et al.* (19) identified and sequenced the 1.3-kb *Eco*RI fragment as part of the human plasminogen gene; it exactly matches the D3 fragment that we cloned and sequenced. Ichinose (22) has more recently reported partial sequence of three highly similar 1.3-kb *Eco*RI fragments, which he designated plasminogen and plasminogen-related genes PRG-A and PRG-B. Our D1 and D2 fragments closely resemble the 400 nucleotides of PRG-A and PRG-B presented by Ichinose (two differences with each) and might comprise allelic variants of those genes. (The three 1.3-kb D fragments contain about 4% overall nucleotide differences.) It is possible that other plasminogen-related genes exist in the "plasminogen-like" collection of 1.3-kb *Eco*RI fragments. Interestingly, it would be a mistake to presume these other members are merely "pseudogenes." Weissbach and Treadwell (31) identified a gene product in cancer cells that encodes a polypeptide of 8800 Da that resembles only the preactivation peptide domain of plasminogen and whose sequence matches that of PRG-B. We have preliminary evidence that points to transcribed products corresponding to the C fragment as well as the apo(a) gene fragment A. [Our C fragment corresponds to the sequence called apo(a) II by Ichinose (22).] Despite the existence of the 5′ homologues, there is only one active apo(a) gene. This is inferred from the highly informative size polymorphism of the apo(a) protein combined with the extensive record of protein immunoblotting results showing that individuals contain two or fewer distinctive apo(a) isoforms (e.g., refs. 1–3, 8, 9, 11, 13, and 14). The 5′ flanking region and exon one of the human apo(a) gene are contained in the 4.8-kb *Eco*RI "A" fragment, which matches the reported sequences of the 400-bp apo(a)I genomic fragment (22) and (with one mismatch) apo(a) cDNA (7).

The 1.4 kb of 5′ flanking genomic DNA of the apo(a) gene fragment A [as well as the apo(a)-related B, C, and plasminogen-like D fragments] contain sufficient promoter/enhancer elements to drive transcription of the luciferase gene when transfected into HepG2 cells. This raises the possibility that unanticipated gene products may be found that correspond to the other fragments, as was the case for the PRG-B sequence (31). The apo(a) gene fragment contains recognizable consensus sequences for nonspecific promoter elements such as TATA and CAAT boxes and for the hepatocyte transcription elements HNF-α1, CEBP, and LF-A1 (32–34). In addition,

sequence motifs were found that correspond to interleukin 6-responsive elements (35, 36). This is consistent with the finding that apo(a) can respond to trauma as an acute-phase protein (37) and our finding that interleukin 6 leads to an ≈5-fold enhancement of reporter gene activity in the transfected HepG2 cells (unpublished data). Evaluation of the role of these sequence elements in the apo(a) gene requires *in vitro* mutagenesis and DNA binding protein studies. Other regions of the apo(a) gene also remain to be tested for the presence of control elements.

Two of the hallmarks of apo(a) gene activity are the striking inter-individual variation in plasma concentration (roughly 1000-fold in the population), contrasted with its relative insensitivity to age, diet, and environmental influences in each given person. Hence we hypothesize that a significant cause of the variation in apo(a) concentration is inherited differences in cis-acting elements of transcriptional control, perhaps of the constitutive or liver-specific variety. Preliminary evidence supporting this may be seen in the different *in vitro* transcription activities we found between the apo(a) genomic fragments isolated from two individuals with high and low plasma levels of apo(a). We expect that other individual differences in the 5′ flanking and other regions of the apo(a) gene will contribute to the inherited variation in apo(a) plasma levels. The isolation and identification of the 5′ regions of the apo(a), plasminogen, and homologous genes provide the tools to begin to study the regulation of these genes, which play important roles in the normal and pathologic status of blood vessels.

We thank Drs. Hiroki Kurihara and George Koike for their kind assistance in this project, which was supported by grants from the Cigarette and Tobacco Surtax Fund of the State of California to R.M.L.; the American Heart Association, California Affiliate, to D.P.W.; the Bugher Foundation, American Heart Association, to J.G.C.; the Fogarty International Center to G.E.L.; and the National Institutes of Health to A.C.L. and B.R.Z.

1. Utermann, G. (1989) *Science* **246,** 904–910.
2. Scanu, A. M. & Fless, G. M. (1990) *J. Clin. Invest.* **85,** 1709–1715.
3. Scanu, A. M., Lawn, R. M. & Berg, K. (1991) *Ann. Intern. Med.* **115,** 209–218.
4. Rhoads, G., Dahlen, G., Berg, K., Morton, N. & Dannenberg, A. (1986) *J. Am. Med. Assoc.* **156,** 2540–2544.
5. Rosengren, A., Wilhelmsen, L., Eriksson, E., Risberg, B. & Wedel, H. (1990) *Br. Med. J.* **301,** 1248–1251.
6. Boerwinkle, E., Leffert, C. C., Jingping, L., Lackner, C., Chiesa, G. & Hobbs, H. H. (1992) *J. Clin. Invest.* **90,** 52–60.
7. McLean, J., Tomlinson, J., Kuang, W., Eaton, D., Chen, E., Fless, G., Scanu, A. & Lawn, R. (1987) *Nature (London)* **330,** 132–137.
8. Utermann, G., Kraft, H. G., Menzel, H., Hopferwieser, T. & Seitz, C. (1988) *Hum. Genet.* **7,** 41–46.
9. Lackner, C., Boerwinkle, E., Leffert, C., Rahmig, T. & Hobbs, H. H. (1991) *J. Clin. Invest.* **87,** 2153–2161.
10. Krempler, F., Kostner, G. M., Bolzano, K. & Sandhofer, F. (1980) *J. Clin. Invest.* **65,** 1483–1490.
11. Knight, B. B., Perombelon, Y. F. N., Soutar, A. K., Wade, D. P. & Seed, M. (1991) *Atherosclerosis* **87,** 227–237.
12. Gavish, D., Azrolan, N. & Breslow, J. (1989) *J. Clin. Invest.* **84,** 2021–2027.
13. Sandholzer, C., Saha, N., Kark, J., Rees, A., Jaross, W., Dieplinger, H., Hopplichler, F., Boerwinkle, E. & Utermann, G. (1992) *Arterioscler. Thromb.* **12,** 1214–1226.
14. Koschinsky, M. L., Beisiegel, U., Henne-Bruns, D., Eaton, D. L. & Lawn, R. M. (1990) *Biochemistry* **29,** 644–646.
15. Wade, D. P., Knight, B. L., Spengel, K. H. & Soutar, A. K. (1991) *Atherosclerosis* **91,** 63–72.
16. Azrolan, N., Gavish, D. & Breslow, J. (1991) *J. Biol. Chem.* **266,** 13866–13872.
17. Sambrook, J., Fritsch, E. & Maniatis, T. (1986) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
18. Chirgwin, J. M., Pryzbyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
19. Peterson, T. E., Martzen, M. R., Ichinose, A. & Davie, E. W. (1990) *J. Biol. Chem.* **265,** 1604–1611.
20. Malinowski, D. P., Sadler, J. E. & Davie, E. W. (1984) *Biochemistry* **23,** 4243–4250.
21. Forsgren, M., Raden, B., Isnaelsson, M., Larsson, K. & Heden, L. O. (1987) *FEBS Lett.* **213,** 254–260.
22. Ichinose, A. (1992) *Biochemistry* **31,** 3113–3118.
23. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1991) *Current Protocols in Molecular Bology* (Wiley, New York), pp. 4.8.1–4.8.3.
24. Malgaretti, N., Bruno, L., Pontoglio, M., Candiani, G., Meroni, G., Ottolenghi, S. & Taramelli, R. (1990) *Biochem. Biophys. Res. Commun.* **173,** 1013–1018.
25. Patthay, L. (1985) *Cell* **41,** 657–663.
26. Castellino, F. J. & Beals, J. M. (1987) *J. Mol. Evol.* **26,** 358–369.
27. Laplaud, P. M., Beaubatie, L., Rall, S. J., Luc, G. & Shaboureau, M. (1988) *J. Lipid Res.* **29,** 1157–1170.
28. Murray, J. C., Buetow, K. H., Donovan, M., Hornung, S., Motulsky, A. G., Disteche, C., Dyer, K., Swisshelm, K., Anderson, J., Giblett, E., Sadler, E., Eddy, R. & Shows, T. B. (1987) *Am. J. Hum. Genet.* **40,** 338–350.
29. Frank, S. L., Klisak, I., Sparkes, R. S., Mohandas, T., Tomlinson, J. E., McLean, J. W., Lawn, R. M. & Lusis, A. J. (1988) *Am. J. Hum. Genet.* **79,** 352–356.
30. Drayna, D. T., Hegele, R. A., Haas, P. E., Emi, M., Wu, L. L., Eaton, D. L., Lawn, R. M., Williams, R. R., White, R. L. & Lalouel, J. M. (1988) *Genomics* **3,** 230–236.
31. Weissbach, L. & Treadwell, B. V. (1992) *Biochem. Biophys. Res. Commun.* **186,** 1108–1114.
32. Courtois, G., Morgan, J. G., Campbell, L. A., Fourel, G. & Crabtree, G. R. (1987) *Science* **238,** 688–692.
33. Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J. & McKnight, S. L. (1988) *Genes Dev.* **2,** 789–800.
34. Hardon, E., Frain, M., Paonessa, P. & Cortese, R. (1988) *EMBO J.* **7,** 1711–1716.
35. Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T. & Kishimoto, T. (1990) *EMBO J.* **9,** 1897–1906.
36. Hattori, M., Abraham, L. J., Northemann, W. & Fey, G. H. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2364–2368.
37. Maeda, S., Abe, A., Seishima, M., Makino, K., Noma, A. & Kawade, M. (1989) *Atherosclerosis* **78,** 145–150.