THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 280, No. 31, Issue of August 5, pp. 28215–28220, 2005

Determination of the Functionality of Common *APOA5* Polymorphisms*

Received for publication, February 24, 2005, and in revised form, May 11, 2005
Published, JBC Papers in Press, June 7, 2005, DOI 10.1074/jbc.M502144200

Philippa J. Talmud‡§, Jutta Palmen‡, Wendy Putt‡, Laurence Lins¶, and Steve E. Humphries‡

From the ‡Division of Cardiovascular Genetics, Department of Medicine, Royal Free and University College Medical School, 5 University Street, London WC1E 6JF, United Kingdom and the ¶Centre de Biophysique Moléculaire Numérique, Faculté des Sciences Agronomiques de Gembloux, 5030 Gembloux, Belgium

Common variants of *APOA5* have consistently shown association with differences in plasma triglyceride (TG) levels. These single nucleotide polymorphisms (SNPs) fall into three common haplotypes: *APOA5*1*, with common alleles at all sites; *APOA5*2*, with rare alleles of −1131T→C, −3A→G, 751G→T, and 1891T→C; and *APOA5*3*, distinguished by the c56C→G (S19W). Molecular modeling of the apoAV signal peptide (SP) showed an increased angle of insertion (65°) at the lipid/water interface of Trp-19 SP compared with Ser-19 SP (40°), predicting reduced translocation. This was confirmed by 50% reduction of Trp-19-encoded SP·secretory alkaline phosphatase (SEAP) fusion protein secreted into the medium from HepG2 cells compared with the Ser-19·SEAP fusion protein ($p < 0.002$). Considering *APOA5*2* SNPs, there was no significant difference in the relative luciferase expression in Huh7 cells transiently transfected with a −1131T construct compared with the −1131C (fragments −1177 to −516 or −1177 to −3). Similarly, for the −3A→G in the Kozak sequence, *in vitro* transcription/translation assays and primer extension inhibition assays showed no alternate AUG initiation codon usage, demonstrating that −3A→G did not influence translation efficiency. Although 1891T→C in the 3′-untranslated region disrupts a putative *Oct-1* transcription factor binding site, when inserted 3′ of the luciferase gene the T→C change demonstrated no significant difference in luciferase expression. Thus, association of *APOA5*2* SNPs with TG levels is not due to the individual effects of any of these SNPs, although cooperativity between the SNPs cannot be excluded. Alternatively, the effect on TG levels may reflect the strong linkage disequilibrium with the functional *APOC3* SNPs.

There is strong evidence from clinical studies, epidemiological studies, and animal models (reviewed in Ref. 1) that apoAV plays a key role in triglyceride (TG)[1] metabolism. Transgenic and knock-out mouse models identified an inverse relationship between apoAV and plasma TG levels (2). *APOA5* variants determine between individual differences in plasma TG in all ethnic groups studied to date (3, 4). The role of apoAV in TG metabolism is further supported by the recently identified rare mutation Q145X, introducing a premature termination of apoAV and resulting in apoAV deficiency and severe hypertriglyceridemia (5). However the exact function(s) of apoAV remain to be clarified. *APOA5* expression is up-regulated in liver regeneration after rat hepatectomy (6), suggesting that it acts as a break on very low density lipoprotein assembly (7). ApoAV is found preferentially on high density lipoprotein but is thought to transfer to very low density lipoprotein in the postprandial state (8) and has been shown to activate lipoprotein lipase *in vitro* and *in vivo* (9).

The *APOA5* gene itself is relatively polymorphic. The five most common *APOA5* haplotypes can be defined by seven SNPs (10). Two haplotypes, *APOA5*2* and *APOA5*3*, are of particular interest as they show association with raised plasma TG levels (4). *APOA5*1* is the wild type haplotype defined by the common alleles at the seven sites. This is found at a frequency of 69% in Caucasian populations (10). Haplotype *APOA5*2* is defined by the rare alleles at four sites (10), −1131T→C (SNP3 (2)), −3A→G, 715G→T (SNP 2, previously referred to as IVS3 + 476G→T (2)), and 1891T→C (SNP 1, previously referred to as c1259T→C (2)) in the 3′-UTR of the gene. The third haplotype, *APOA5*3*, is distinguished from *APOA5*1* by the rare allele of the cSNP 56C→G, which results in the substitution of Trp for Ser at residue 19 (S19W; SNP5 (2)). Haplotypes *APOA5*2* and *APOA5*3* occur at frequencies of 4% in a Caucasian study (11). The rare alleles of 1764C→T and V153M, the two remaining SNPs, both on chromosomes defined by the other common alleles, represent the other two *APOA5* haplotypes. Family studies suggest that V153M is nonfunctional because it shows no co-segregation with hypertriglyceridemia in a carrier family (4), whereas there is no information about the association of 1764C→T with plasma lipids. We have therefore concentrated our functional studies on SNPs that define haplotypes *APOA5*2* and -**3*.

The proximity of the *APOA5* to *APOA4/APOC3/APOA1* in the same gene cluster on chromosome 11q23 raised the question of whether the TG-raising effects of *APOA5* SNPs simply reflected linkage disequilibrium (LD) with functional variants in the other TG-raising gene, *APOC3*, or whether these were independent functional effects (3). Haplotype analysis of the *APOA5/A4/C3* gene cluster identified that whereas the S19W showed association with plasma TG levels, independent of *APOC3*, the −1131T→C is in strong LD with the *APOC3* −482C→T, and thus associations of haplotype *APOA5*2* may indeed be due to the effects of −482C→T, which disrupts the normal insulin responsiveness of apoCIII. No detailed functional studies have been reported on these two TG-raising *APOA5* haplotypes, and thus the purpose of this study was to investigate the basis for their effects and to determine which of these variant sites are functional.

* This work was supported by British Heart Foundation Grants PG/2000/057, PG/2000/134, RG2000/015, and PG/04/110/17827. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. Tel.: 44-207679-6968; Fax: 44-207679-6212; E-mail: p.talmud@ucl.ac.uk.

[1] The abbreviations used are: TG, triglyceride(s); SNP, single nucleotide polymorphism; UTR, untranslated region; LD, linkage disequilibrium; SP, signal peptide; SEAP, secretory alkaline phosphatase; nt, nucleotide(s).

MATERIALS AND METHODS

Molecular Modeling of apoAV Signal Peptide Variant

The Ser-19 (wild type) and variant Trp-19 23-amino acid signal peptides (SP) of apoAV were three-dimensionally constructed as α-helices using Hyperchem 5.0 (Hypercube, Inc.). Their conformations were minimized by the Polak-Ribiere algorithm in an AMBER force field. The two peptides were analyzed by the IMPALA algorithm, described elsewhere (12). This enables analysis of the insertion of a peptide in a modeled membrane using simple restraint functions designed to mimic the membrane properties. The position of the peptide in the modeled membrane is minimized by a Monte Carlo procedure.

*Functional Studies of S19W SNP Defining Haplotype APOA5*3*

Replacement of the Secreted Alkaline Phosphatase Signal Peptide with the ApoAV Signal Peptide Variants, Ser-19 and Trp-19

To test the effect of the apoAV signal peptide variant S19W, *APOA5* signal peptide sequences with either a C or G at position 56 were used to replace the endogenous signal peptide sequence of the secreted alkaline phosphatase (SEAP) in the Great EscAPe SEAP Reporter System 3 vector (BD Biosciences) by PCR concatenation following the method of Tuohy and Groden (13). The forward primer consisted of a SEAP sequence 5′ to its signal peptide followed by the *APOA5* signal peptide sequence: internal forward (IntF) 5′-ATCGCCCACCATGCAGATAAT-GGCAAGCATGGCTGCCGTGCTCACCTGGGCTCTGGCTCTTGTTT-CAGCG-3′. The reverse primer contained an overlap with a post-signal peptide SEAP sequence and the 3′ end of the *APOA5* signal sequence: internal reverse (IntR) 5′-GCAACTGGGATGATTGCCTGGGTGGC-C<u>G</u>AAAACGCTGAAAGAACAGCCAGAGCCA-3′ (the base 28 from the end of the primer (underlined) was either a G or C depending on whether the wild type (Ser-19) or variant (Trp-19) product was being amplified). External primers external forward (ExtF) ATCGC<u>GAAT-TC</u>GCCCACCATGC and external reverse (ExtR) 5′-CTCAAG<u>CCAAT-GGTCTGG</u>AAGTTCG-′ with EcoRI and BstXI sites (underlined), respectively, overlapped the IntF oligo and a naturally occurring BstXI site in the SEAP vector to allow generation of a 337-bp PCR product. Equal amounts (100 ng) of SEAP control vector and the *APOA5*/pGEM7 plasmid were mixed and concatenated by PCR in a 100-μl reaction mix containing 1× cloned *Pfu* buffer (Stratagene), 25 pmol of each primer, IntF, IntR, ExtF, and ExtR, 0.2 mM dNTPs, and 5 units of *Pfu Taq* polymerase (Stratagene). The cycling conditions were 94 °C for 2 min, five cycles of 94 °C for 45 s, 45 °C for 45 s, and 72 °C for 2 min, 30 cycles of 94 °C for 45 s, 48 °C for 45 s, and 72 °C for 2 min, and 72 °C for 5 min to produce a PCR product size of 337 bp. The SEAP control vector was prepared by excising the EcoRI/BstXI fragment prior to subcloning the similarly digested *APOA5* PCR product into its place. The final PCR products were checked by sequencing using a Big Dye terminator kit (Applera) and an ABI 377 automated sequencer.

Assaying Secreted Alkaline Phosphatase Expression

HepG2 cells were maintained in RPMI + L-glutamine supplemented with 10% fetal bovine serum. SEAP was assayed after transfection of the various constructs into HepG2 cells in OptiMEM (Invitrogen), using FuGENE 6 (Roche Applied Science) following the manufacturer's protocol for the 96-well format. The Basic SEAP (lacking promoter and enhancer sequences) and the unmodified Control SEAP vector (with promoter and enhancer) acted as negative and positive controls, respectively. Co-transfection with β-galactosidase acted as a control for transfection normalization. The medium from the transfected cells was collected after a 48-h incubation. The cells were fixed and stained for β-galactosidase, and secreted alkaline phosphatase was detected using 4-methylumbelliferyl phosphate as substrate according to the Great EscAPe SEAP Reporter System 3 protocol (BD Biosciences). The excitation and emission peaks of 4-methylumbelliferyl phosphate fluorescence, 360 nm and 449 nm, respectively, were measured on a Tropix TR717 microplate luminometer (Applied Biosystems) to estimate quantitatively the amount of fusion protein secreted from the cells. Five repeat transfections were carried out, and each sample was replicated five times.

*Functional Studies of Variants Comprising haplotype APOA5*2*

Promoter −1131T→C

To test the function of the −1131T→C promoter variant, four different fragments were amplified by PCR from the DNA of individuals with known −1131T→C and −3G→A genotype. Two short fragments, which included only the −1131 site and were referred to as −1131T and −1131C, ranged from position −1177 to −512 using the original numbering of Olivier *et al.* (10). A second set of longer fragments, which incorporated both the −1131 and −3 sites with the naturally occurring combinations and referred to as −1131T/−3A and −1131C/−3G, ranged from position −1177 to −3. Oligonucleotides were designed with KpnI sites at their 5′ end of the forward oligo and BglII at the 5′ end of the reverse oligo. In addition CGG and GA tails were introduced at 5′ ends of forward and reverse oligos. Oligonucleotides for the short fragments were as follows: forward primer, 5′-CGGGGTACCAGAGGCCCT-GCGAGTGGAGTT-3′; reverse primer, 5′-GAAGATCTTGCTCACCT-GCTCACGTCT GG-3′.

For the long fragment the same forward oligonucleotide was used. For the reverse primers the oligonucleotides were: reverse primer −1131T/−3A, 5′-GAAGATCTTCTGCTCTGAGAAGACAG GTGG-3′; reverse primer −1131C/−3G, 5′-GAAGATCTCCTGCTCTGAGAAGA-CAG GTGG-3′.

PCR products were TA-cloned into pGEM-T vector system (TM042) following the manufacturer's conditions (Promega) and subsequently cloned into the KpnI/BglII polylinker site of the pGL3 Basic luciferase reporter vector (TM033). Minipreps were purified using the QIAprep Spin miniprep kit (Qiagen) and sequenced with Applera BigDye Terminator v3.1 sequencing kit on an ABI 377 automated sequencer. Maxipreps were prepared with the GenElute HP plasmid Maxiprep kit (Sigma). Two different DNA preparations were used.

Both Huh7 and HepG2 cells were used for transient transfection. Results were essentially the same; results from Huh7 cells are shown here. Huh7 cells in 96-well format (4×10^5 cells/well) were transiently transfected with 200 ng of the appropriate plasmid or control vector, 2 ng of pRL-TK co-transfector, and 0.5 μl of Lipofectamine 2000, and the cells were lysed 48 h after transfection. Luciferase activity was assayed using the Dual luciferase reporter assay system (TM040, Promega). Results represent the mean from nine repeat experiments, each performed in sets of eight repeat wells.

Kozak Sequence −3A→G

Two different assays were performed to test the functionality of the −3A→G variant.

In Vitro Transcription and Translation Assay—Two *APOA5* cDNA clones (5′-UTR and coding sequence only), a kind gift from Len Pennacchio, were used: 3A/56C and −3G/56C. The cDNAs were subcloned into pGEM 7Zf driven from a T7 promoter (Promega) and confirmed by sequencing. The transcription/translation experiments were performed by means of the TNT quick-coupled transcription/translation system (Promega). The plasmids were transcribed *in vitro* and translated by using the TNT-coupled reticulocyte lysate system (Promega TNT Quick transcription/translation system) following the manufacturer's protocol and using FluoroTect Green (Promega FluoroTect™ Green $_{LYS}$ tRNA) to label newly synthesized protein. Size and quantification of the proteins were checked by SDS-PAGE, read on a Typhoon 8600 (Molecular Dynamics) at 532 nm excitation, and analyzed with ImageQuant software. Three separate DNA preparations of each clone were used, and samples were duplicated for each experiment.

Primer Extension Inhibition (Toe Printing) Assay—To study the initiation AUG of protein synthesis, a primer extension inhibition assay (14) in rabbit reticulocyte lysates was used. The same constructs were used as in the transcription/translation assay. Plasmid DNA, linearized by digestion with EcoRI, provided the template for transcription by T7 RNA polymerase using the Riboprobe System-T7 (Promega) and Ribo m^7G cap analog (Promega). Capped mRNAs were purified by extraction with phenol, concentrated by ethanol precipitation, resuspended in water, and stored at −70 °C. An oligonucleotide, AGAGGCCT-CAGCTTTTCCAGG, labeled at the 5′ end with 6-carboxyfluorescein (6-FAM), was used to prime the reverse transcriptase step. The fluorescently labeled primer was pre-annealed to the RNA by heating for 1 min at 65 °C followed by incubation at 37 °C for 8 min in 40 mM Tris-HCl, pH 7.5, and 0.2 mM EDTA. The primer-RNA complexes were then held on ice for ~15 min, whereas the reticulocyte reaction mixtures were assembled. Ribosome binding reactions were carried out using a micrococcal nuclease-treated rabbit reticulocyte lysate (Promega). The reaction mixes contained 45% reticulocyte lysate, 90 μg/ml cycloheximide, 200 μM sparsomycin, 2 mM magnesium $(CH_3COO)_2$, and 100 mM potassium acetate. For the assay, 25-μl aliquots of this mixture containing 2 μl of mRNA/primer (0.1–0.2 μg of mRNA) were incubated at 25 °C for 6 min and then diluted with 20 volumes of cold buffer, 50 mM Tris-HCl, pH 7.5, 40 mM KCl, 6 mM $MgCl_2$, 5 mM dithiothreitol, 110 μg/ml cycloheximide, and each of the four dNTPs at 575 μM. Primer extension was initiated by adding 2 units/μl Superscript II reverse transcriptase to the mix and incubated at 25 °C for 10 min. Reactions were stopped by extraction with phenol and precipitation with ethanol before resuspension in 10 μl of water. 2-μl aliquots, together with GeneScan-500 ROX (dichlororhodamine acceptor dye) size standards (Applera), were run on an ABI 377 automated sequencer and analyzed using GeneScan software. Experiments were repeated twice, each time in quadruplicate.

3′-UTR 1891T→C

To test the function of the 1891T→C variant in the 3′-UTR, an 816-bp fragment was amplified by PCR from the DNA sample from an individual with haplotype *APOA5*1* from +1752 to +2568 with a T at position 1891. Oligonucleotides were designed with XbaI sites at the 5′ end of the forward oligo and HpaI at the 5′ end of the reverse oligo. In addition, GG and CG tails were introduced at the 5′ ends of forward and reverse oligos. Oligonucleotides for the short fragments were as follows: forward primer, 5′-GCTCTAGA GGCCCATTCCCAGCTCCTTGT-3′; reverse primer, 5′-GGGTTAACTGTGCTTTGGGGATAGTGGTGAGG-3′.

PCR products were TA-cloned into pGEM-T vector system (TM042) following the manufacturer's conditions (Promega). Site-directed mutagenesis was used to introduce the T→C at position 1891 using the Stratagene QuikChange site-directed mutagenesis kit (Stratagene) and confirmed by sequencing. Subcloning into the XbaI/HpaI polylinker site 3′ of either the pGL3 control or promoter Luciferase reporter vector (TM033) followed. Minipreps were purified using the QIAprep Spin miniprep kit (Qiagen) and sequenced with Applera BigDye Terminator v3.1 sequencing kit on an ABI 377 automated sequencer. Maxipreps were prepared with the GenElute HP plasmid Maxiprep kit (Sigma).

Huh7 cell transfection and luciferase activity measures were carried out as described above for the −1131T→C variant.

RESULTS

*Functional Predictions and Effects of Haplotype APOA5*3*

Molecular Modeling of the apoAV Signal Peptide S19W Variant—The first 23 N-terminal residues of apoAV form a signal peptide. We had previously shown that signal peptides insert obliquely into the lipid membrane (15, 16). This tilted orientation was important to the translocation activity of the signal peptide, with an angle of insertion of 45° giving the highest activity (15). The most stable position in a modeled membrane for the Ser-19 SP and the Trp-19 SP are shown in Fig. 1. The Ser-19 SP assumes an angle of insertion relative to the phospholipids bilayer of 40°, with its mass center at 10 Å from the bilayer center. For the Trp-19 SP, the best position relative to the bilayer is at an angle of 65°, with its mass center at 5 Å. For comparison, modeling of the endogenous secretory alkaline phosphatase SP shows an angle of insertion of 40° relative to the bilayer plane with a mass center at 9 Å from the bilayer center (data not shown). This increase of the angle of insertion from 40° to 65° for Ser-19 compared with Trp-19 is predicted to lead to reduced "activity," *i.e.* translocation across the endoplasmic reticulum. These predictions also suggest that the apoAV Ser-19 SP should work as efficiently as the endogenous SP when fused to SEAP.

Influence of the APOA5 Signal Peptide Variant S19W on Secretory Alkaline Phosphatase Secretion—The effect of the substitution of the SEAP signal sequence by the *APOA5* signal sequence, encoding either Ser or Trp at residue 19 and creating two different apoAV·SEAP fusion proteins, was examined. The secretion of alkaline phosphatase from HepG2 cells into the medium was then assayed quantitatively. As presented in Fig. 2, compared with control SEAP with its endogenous signal peptide, the apoAV wild type Ser-19 signal peptide acted effectively as a signal sequence, as predicted by the molecular modeling. However, substitution of Ser-19 by Trp-19 reduced the amount of secreted SEAP by 49% ($p = 0.002$), proving that this SNP is indeed functional. According to the manufacturer's protocol, the SEAP in the medium is directly proportional to the intracellular mRNA and protein. For this reason the intracellular concentration of the fusion proteins was not measured.

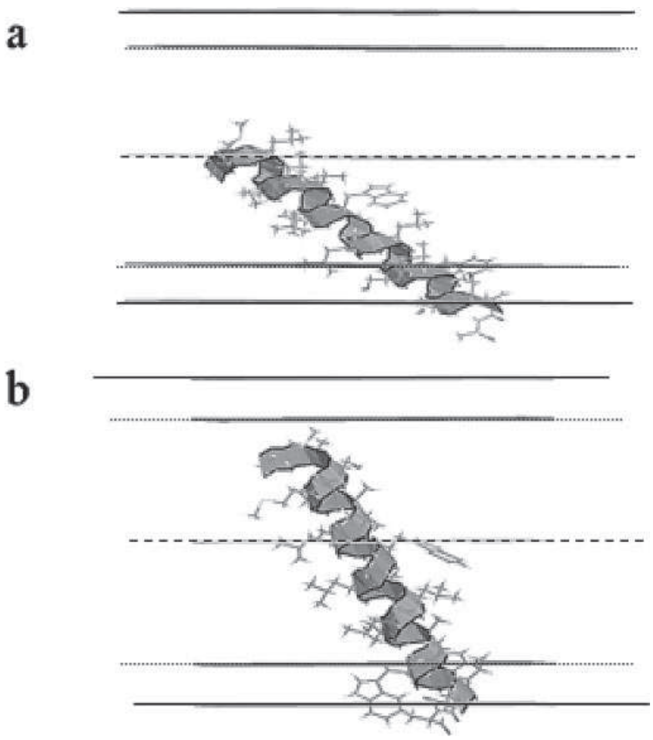

FIG. 1. **Molecular modeling of Ser-19 and Trp-19 signal peptides inserted in a simplified lipid bilayer.** Best position for the Ser-19 (*a*) and Trp-19 (*b*) signal peptides after IMPALA minimization (12) is shown. The *dashed line* represents the phospholipid head group/water interface (placed at 18 Å from the bilayer center); the *dotted line* represents the phospholipid head group/phospholipid acyl chain interface (at 13 Å from the bilayer center); and the *solid line* represents the bilayer center (at the origin ($z = 0$)). The bilayer is symmetric. The peptides are represented as *ribbons*.

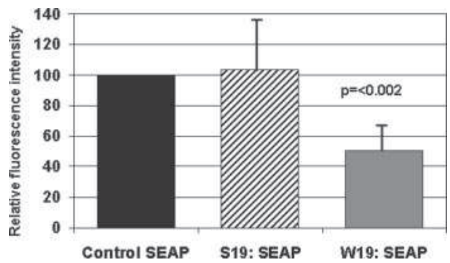

FIG. 2. **Quantitative estimation of secretion of the apoAV signal peptide·SEAP fusion proteins from HepG2 cells.** HepG2 cells were transiently transfected with the wild type SEAP construct with its endogenous signal peptide, the *APOA5* Ser-19·SEAP or *APOA5* Trp-19·SEAP constructs. Co-transfection with β-galactosidase acted as a control for transfection normalization, and the SEAP basic vector acted as the negative control. Wild type SEAP was set at 100%. Comparison of the amount of secreted apoAV Ser-19·SEAP and apoAV Trp-19·SEAP into the medium shows a statistically significant difference ($p = 0.002$). Results represent five separate transfections, and each sample was replicated five times. *Bars* = S.D.

*Functional Effects of Haplotype APOA5*2*

Haplotype *APOA5*2* is defined by four variant sites all in complete LD, namely −1131T→C,-3A→G, 715G→T, and 1891T→C. Of these variants, 715G→T, 476 bp into intron 3, is unlikely to be functional and was not examined further. The potential functional effects of the other three variants were tested.

Promoter −1131T→C—This variant within the promoter of *APOA5* could alter transcription, and in order to study this further, its effect on the expression of luciferase, acting as a reporter gene, was examined. To test whether −1131T→C was acting independently or could be influenced by the −3A→G

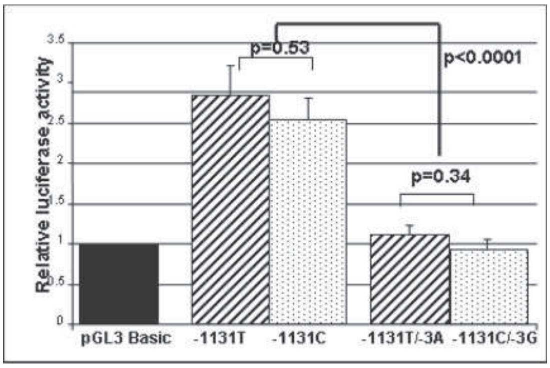

FIG. 3. **Effect of the promoter variant −1131T→C on relative luciferase expression.** pGL3 Basic vector set as 1. −1131T and −1131C represent luciferase expression from constructs with the short fragments, −1177 to −516; −1131T/−3A and −1131C/−3G represent results from constructs with the long fragments, −1171 to −3. Data were corrected for transfection efficiency. Results represent data from nine repeat experiments, each with eight repeats, and two different DNA preparations. *Bars* = S.D.

change (with which it shows complete LD), four different constructs were made, two shorter fragments (from −1177 to −516) that included only the −1131 site, and two larger constructs that included in addition the −3 site (from −1177 to −3) in the naturally occurring combinations (−1131T/−3A and −1131C/−3G). The effects of these four constructs on luciferase expression in the luciferase dual assay are presented in Fig. 3. Results from nine separate experiments on two different DNA preparations (each with eight repeats) showed no significant difference between the −1131C or −1131T short constructs on luciferase expression ($p = 0.53$), although these constructs showed more than double the effect on luciferase activity compared with the pGL3 Basic vector. Similarly there was no statistically significant difference between the −1131T/−3A and −1131C/−3G constructs ($p = 0.34$), although both constructs showed significantly lower luciferase activity compared with the short constructs ($p < 0.0001$).

Kozak Sequence −3A→G—Two different assays were used to test the function of this variant. The *in vitro* transcription/translation assay was used to investigate whether the −3A→G in the Kozak sequence affected the translation efficiency of −3G and −3A cDNAs. Evaluation was carried out by lysine fluorescence (a representative gel is shown in Fig. 4*a*), with the major band representing the translated apoAV proteins. Quantification of these results from three repeat experiments (Fig. 4*b*) shows that there was no significant difference ($p = 0.59$) in the fluorescence emission between the two constructs, establishing that the site at −3 was not influencing translation efficiency.

This was confirmed by the complementary primer extension inhibition "toe printing" assay, which examines codon usage. The premise is that the −3A→G change in the Kozak sequence could result in leaky translation, such that translation might start at an alternative AUG initiation site, because of poor discrimination. ApoAV has AUG codons at residues 1 and 4. The essence of this assay is that the interaction of ribosomes with mRNA can be limited to the initiation step by including antibiotics that inhibit elongation (14, 17). A 5′ -fluorescently labeled oligo, 230 nt 3′ of the residue 1 start site and 221 nt 3′ of residue 4, was used to prime reverse transcriptase to the position of the bound ribosome. Representative electropherograms from the two constructs −3A or −3G are shown in Fig. 5. As expected, the full-length 5′ end product of 285 bp was visible. A 253-bp fragment was present for both constructs and is likely to represent an echo band (14). As reported by Kozak (14), the 80 S ribosome bound to the AUG will protect 15–16 nt

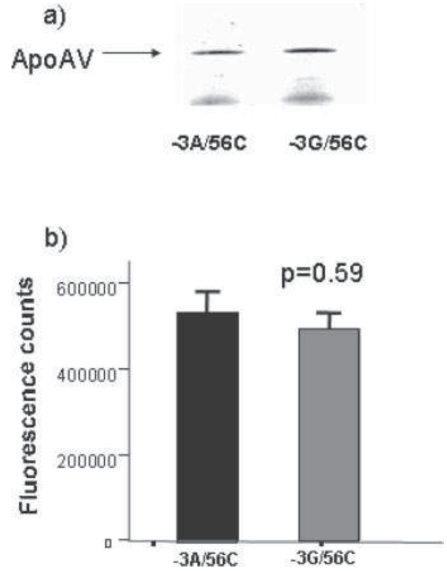

FIG. 4. ***In vitro* translation/transfection assay on the Kozak sequence variant −3A→G.** *a*, representative gel showing a comparison of the translated apoAV from a −3A/56C or −3G/56C construct. *b*, quantification of the fluorescence from three repeat experiment, each in duplicate, from −3A/56C or −3G/56C cDNAs ($p = 0.059$).

3′ of the codon, and thus the expected bands, if both AUGs are used, should be 214 or 205 nt long. The peak at 209 nt probably represents the 80 S ribosome blocked codon 1 fragment, which is slightly shorter than expected, representing 21 nt blocked by the 80 S ribosome. This might be because of distortion by the fluorescent label. A small peak at 230 nt was seen with both constructs, which would be the expected size fragment corresponding to codon 1 usage. No peak at 221 nt (which would represent codon 4 usage) was seen. Importantly, no second additional peak representing the 9-nt shorter fragment was evident in the −3G construct. These results thus confirm results from the transcription/translation assay by demonstrating that only the AUG at residue 1 is used and the −3A→G change does not result in leaky translation.

3′-UTR 1891T→C—The function of the T→C change in the 3′-UTR of *APOA5*, acting as a 3′-UTR for the luciferase gene, was tested in two different vectors, pGL3 control vector (which in addition to the SV40 promoter has the SV40 enhancer) and pGL3 promoter vector (which lacks the SV40 enhancer). There was no significant difference in the effect of the T or C construct on luciferase expression in either vector system (Fig. 6).

DISCUSSION

The novel aspect of this study is the resolution of the functionality of *APOA5* SNPs by demonstrating the altered signal peptide function of the *APOA5*3* S19W and the lack of functional effects of any of the three *APOA5*2* SNPs tested individually.

*Haplotype APOA5*3*—56C→G, resulting in S19W, introduces an aromatic amino acid, Trp, in exchange for a polar Ser at position −5 from the cleavage site of the apoAV signal peptide. Although the sequences of signal peptides are not highly conserved, they do observe certain constraints, with positively charged residues at the N terminus, a hydrophobic core, and neutral but polar residues at the C terminus (18). Amino acids at −3 and −1 relative to the cleavage site must be small and neutral for cleavage to occur (19). The rigidity of this requirement is borne out by the amino acid substitutions of Ala by Thr at the −1 cleavage site of vasopressin-neurophysin II (20), a Gly to Ala change at −1 in collagen X (21), and a Pro to Ser substitution at −3 in the preproparathyroid hormone (22), which cause severe genetic disorders. However, amino acid

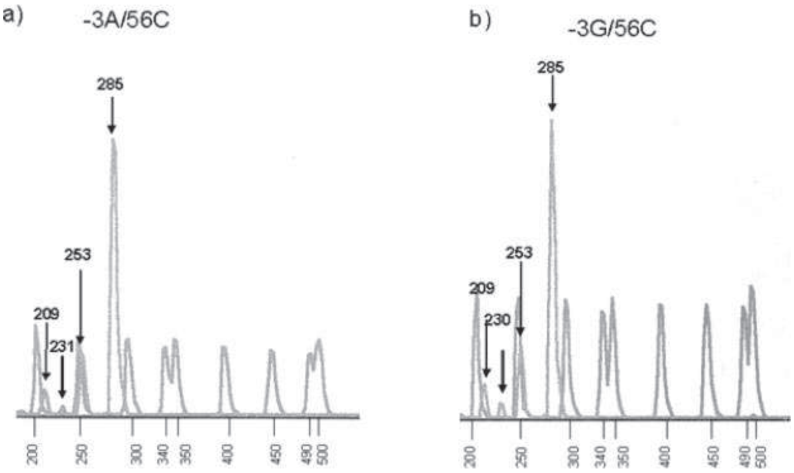

FIG. 5. **Toe printing assay examining the fidelity of primer extension inhibition of −3A/56C (*A*) and −3G/56C cDNA (*B*).** A representative electropherogram shows peaks labeled with the peak size, which represents the fragment sizes from the primer extension, and marker peaks representing the GeneScan-500 ROX size standards.

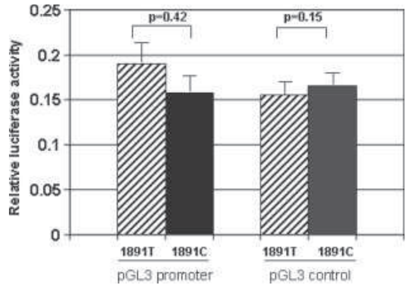

FIG. 6. **Effect of the 3′-UTR 1891T→C on luciferase activity.** Relative luciferase expression from pGL3 control and pGL3 promoter constructs of *APOA5* 3′-UTR 1891T→C acting as a 3′-UTR to the luciferase reporter gene is shown. There was no statistically significant difference between the 1891T and 1891C constructs using either vector. Results represent four separate experiments each with eight repeat runs.

changes at less critical residues within the signal sequence are not associated with such extreme phenotypes. Alteration to the hydrophobic core of the ApoB signal peptide, due to deletion of a Leu-Ala-Leu triplet, reduces the total hydrophobic potential of the signal peptide (15) and is predicted to decrease translocation across the endoplasmic reticulum. This has been confirmed *in vitro* in both heterologous and homologous systems (23, 24). The biological consequence is that carriers of the *APOB* deletion allele have modestly lower plasma TG than those homozygous for the insertion allele (25). The *APOA5* S19W change has a similar effect on SP function as the ApoB SP deletion and alters the predicted angle of insertion of the modeled signal peptide to the water/lipid interface from 40° to 65°. The prediction of reduced translocation and therefore secretion of Trp-19 apoAV compared with Ser-19 apoAV was fully supported by the apoAV SP·SEAP fusion experiments. This fusion protein showed almost 50% lower SEAP in the medium of HepG2 cells compared with the Ser-19·SEAP fusion protein. One caveat is that the intracellular SEAP concentration was not measured; however, a difference between intracellular Ser-19·SEAP fusion protein and Trp-19·SEAP fusion protein concentrations would be difficult to interpret because it might not reflect true intracellular levels because of potential differences in degradation of the two proteins. This reduced secretion would be predicted to result lower plasma apoAV in Trp-19 subjects, although this has not been measured. Because mouse models show an inverse relationship between plasma apoAV and plasma TG levels (2), reduced apoAV should thus result in higher plasma TG levels. We previously reported that men homozygous for the Trp-19 allele had 50% higher plasma TG levels than Ser-19 homozygotes (3), which correlates well with the prediction from the functional studies.

*Haplotype APOA5*2*—In the luciferase reporter gene assay, neither the short constructs, with either the −1131T or −1131C, nor the longer constructs, which incorporated both −1131T/−3A or −1131C/−3G, showed significant differences in luciferase expression levels. We therefore conclude that this base change within the promoter does not affect promoter activity. This is not unexpected because no obvious transcription binding site could be identified spanning or in close proximity to −1131 (3). Although no experimental data were shown, Prieur *et al.* (26) came to the same conclusion about the lack of functionality of this site, using a fragment essentially the same as that in our shorter construct. However, there was a significant difference between the promoter activity of the short *versus* the long constructs. Both short constructs (−1131T or C) resulted in more than 100% higher luciferase expression compared with the larger constructs ($p < 0.0001$). This suggests a possible repressor element between −516 and −3, a region not previously studied. A detailed analysis of the potential transcription factors between −516 and −3 using TRANSPLORER^{TM}, a transcription factor search engine, revealed the presence of several potential nuclear factor binding sites, the most interesting being two potential *Usf-2* sites at −467 to −454 and −23 to −16, two *CEBP* sites at −410 to −397 and −280 to −276, and an *NFkB* site at −77 to −68. ApoAV has been shown to be a positive acute phase protein (27), and the presence of the *NFkB* site suggests a role for this site in the acute phase response.

The Kozak sequence, from −6 to −1 preceding the initiating AUG codon, represents a highly conserved motif in eukaryotes which regulates ribosomal translation (14, 17). The transcription/translation assay used to examine the change within the Kozak sequence, −3A→G, had previously been used to differentiate between the translation efficiency of a C→T change at −1 of the Kozak sequence in annexin V (28). However, the *APOA5* −3A→G change did not affect the translation efficiency of apoAV. The primer extension inhibition assay was used to validate these results and demonstrated that leaky translation, which would have resulted in translation initiation from the +4 AUG rather than +1 AUG of *APOA5*, did not occur. 97% of vertebrates have a purine at position −3 (29), and thus the change from A→G is conservative and unlikely to affect translation efficiency. These two experiments therefore confirmed the lack of functionality of this change within the Kozak sequence

Initial analysis using TRANSPLORER^{TM} to identify potential transcription factor binding sites in the *APOA5* 3′-UTR identified a putative *Oct-1* site (upper case shows the expected consensus sequence for *Oct-1*; lower-case letters do not fit the consensus sequence), CtCctgCa**T**atCCAG (+1883 to +1897), which would be disrupted by the 1891T→C change. However, using this *APOA5* sequence as a 3′-UTR for luciferase showed that the 1891T→C change did not alter luciferase expression, confirming that this SNP is also not functional. Thus, functional studies of the three *APOA5*2* SNPs (the intron 3 SNP is unlikely to be functional), showed that none of these sites significantly altered transcriptional strength, translation efficiency, or 3′-UTR function.

There are three possible explanations for the association of this haplotype with raised plasma TG levels. First, there could be an additional functional site in *APOA5* that is in LD with these three variant sites. However, extensive resequencing (10, 30) has not identified any other common, potentially functional SNPs that could account for the *APOA5*2* haplotype association. Second, these three SNPs of *APO5*2*, although individually not functional, might act cooperatively and thus affect expression levels. An example of such interaction is seen with the α1-antitrypsin gene, where the 11478G→A (TaqI) SNP in the 3′-UTR disrupts an *Oct-1* site and acts in cooperation with a promoter NF-IL6 site as part of the acute phase response (31). Finally, the strong LD between the −1131T→C and the *APOC3* −482C→T variant, identified in our haplotype analysis (3) and subsequently confirmed by Olivier *et al.* (10), suggests that this *APOC3* SNP, which disrupts the insulin responsiveness of *APOC3* leading to raised TG levels (32, 33), could be the functional change influencing plasma TG levels and that the *APOA5*2* haplotype is marking this functional variant.

Although *in vitro* studies have their limitations, they do clarify which SNP(s) are functional. However, they do not provide insight into the functional role of apoAV in TG metabolism. Results from Q145X carriers, as well as the adenovirally mediated gene transfer of *apoa5* into wild type mice, show that lipoprotein lipase activation by apoAV may play an important part in apoAV function (5, 9) as well as the repression of very low density lipoprotein production rate (9). Further detailed functional studies to clarify these exact mechanisms are eagerly awaited.

Evolutionary Implications of the APOA5 SNPs—What emerges from the analyses of these two *APOA5* TG-raising haplotypes in different ethnic groups are the possible differences in their evolutionary history. The frequency of the *APOA5*2*-defining −1131C allele differs significantly among ethnic groups, ranging from 0.34 to 0.37 in Japanese (34, 35), 0.26 in Chinese (36), 0.12 in African-American (4), 0.078 in Hispanic (4), to 0.06 in Caucasian (3) populations. In the CEPH (Centre d' Etude Polymorphisme Humain) family haplotype analysis, despite the indication that there is a region of low LD between *APOA5* and *APOC3*, there was strong LD between *APOA5*2* SNPs and the *APOC3* SNPs (−482C→T, −455T→C, and 3238G→C), with 85% of *APOA5*2* haplotypes occurring on a *APOC3* 3238C allele (10). From an evolutionary viewpoint this suggests that these *APOA5*2* SNPs occurred on a chromosome carrying the rare alleles of these *APOC3* SNPs, prior to population diversification. The consistent association of the −1131T→C with differences in TG across ethnic groups suggests that LD with *APOC3* would be the same in all ethnic groups. There is however no data to confirm this proposition. The marked frequency of differences between ethnic groups probably reflects population bottlenecks and later expansions.

The minor allele frequency of the S19W ranges from 0.15 in Hispanics (4), 0.071 in African-Americans, to 0.057 in Caucasians (3); however it is almost nonexistent in the Japanese, with a frequency of 0.006 (35). The Trp-19 allele occurs only on a "wild type" allele, and its low frequency in those ethnic groups studied, as well as absence in the Japanese (there is no information about Chinese populations), suggests that it is a more recent variant than the *APOA5*2* SNPs and occurred after the diversification of the ethnic groups.

In conclusion, these studies highlight the complex genotype-phenotype relationship that can occur between genotypes (in this case in different genes) that show strong LD.

REFERENCES

1. van Dijk, K. W., Rensen, P. C., Voshol, P. J., and Havekes, L. M. (2004) *Curr. Opin. Lipidol.* **15,** 239–246
2. Pennacchio, L. A., Olivier, M., Hubacek, J. A., Cohen, J. C., Cox, D. R., Fruchart, J. C., Krauss, R. M., and Rubin, E. M. (2001) *Science* **294,** 169–173
3. Talmud, P. J., Hawe, E., Martin, S., Olivier, M., Miller, G. J., Rubin, E. M., Pennacchio, L. A., and Humphries, S. E. (2002) *Hum. Mol. Genet.* **11,** 3039–3046
4. Pennacchio, L. A., Olivier, M., Hubacek, J., Krauss, R. M., Rubin, E. M., and Cohen, J. C. (2002) *Hum. Mol. Genet.* **11,** 3031–3038
5. Oliva, C. P., Pisciotta, L., Li, V. G., Sambataro, M. P., Cantafora, A., Bellocchio, A., Catapano, A., Tarugi, P., Bertolini, S., and Calandra, S. (2004) *Arterioscler. Thromb. Vasc. Biol.* **25,** 411–417
6. van der Vliet, H. N., Sammels, M. G., Leegwater, A. C., Levels, J. H., Reitsma, P. H., Boers, W., and Chamuleau, R. A. (2001) *J. Biol. Chem.* **276,** 44512–44520
7. Weinberg, R. B., Cook, V. R., Beckstead, J. A., Martin, D. D., Gallagher, J. W., Shelness, G. S., and Ryan, R. O. (2003) *J. Biol. Chem.* **278,** 34438–34444
8. Fruchart-Najib, J., Bauge, E., Niculescu, L. S., Pham, T., Thomas, B., Rommens, C., Majd, Z., Brewer, B., Pennacchio, L. A., and Fruchart, J. C. (2004) *Biochem. Biophys. Res. Commun.* **319,** 397–404
9. Schaap, F. G., Rensen, P. C., Voshol, P. J., Vrins, C., van der Vliet, H. N., Chamuleau, R. A., Havekes, L. M., Groen, A. K., and van Dijk, K. W. (2004) *J. Biol. Chem.* **279,** 27941–27947
10. Olivier, M., Wang, X., Cole, R., Gau, B., Kim, J., Rubin, E. M., and Pennacchio, L. A. (2004) *Genomics* **83,** 912–923
11. Eichenbaum-Voline, S., Olivier, M., Jones, E. L., Naoumova, R. P., Jones, B., Gau, B., Patel, H. N., Seed, M., Betteridge, D. J., Galton, D. J., Rubin, E. M., Scott, J., Shoulders, C. C., and Pennacchio, L. A. (2004) *Arterioscler. Thromb. Vasc. Biol.* **24,** 167–174
12. Ducarme, P., Rahman, M., and Brasseur, R. (1998) *Proteins* **30,** 357–371
13. Tuohy, T. M., and Groden, J. (1998) *Hum. Mutat.* **12,** 122–127
14. Kozak, M. (1998) *Nucleic Acids Res.* **26,** 4853–4859
15. Talmud, P., Lins, L., and Brasseur, R. (1996) *Protein Eng.* **9,** 317–321
16. Brasseur, R. (2000) *Mol. Membr. Biol.* **17,** 31–40
17. Kozak, M. (1997) *EMBO J.* **16,** 2482–2492
18. von Heijne, G. (1985) *J. Mol. Biol.* **184,** 99–105
19. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng.* **10,** 1–6
20. Siggaard, C., Rittig, S., Corydon, T. J., Andreasen, P. H., Jensen, T. G., Andresen, B. S., Robertson, G. L., Gregersen, N., Bolund, L., and Pedersen, E. B. (1999) *J. Clin. Endocrinol. Metab.* **84,** 2933–2941
21. Chan, D., Ho, M. S., and Cheah, K. S. (2001) *J. Biol. Chem.* **276,** 7992–7997
22. Sunthornthepvarakul, T., Churesigaew, S., and Ngowngarmratana, S. (1999) *J. Clin. Endocrinol. Metab.* **84,** 3792–3796
23. Sturley, S. L., Talmud, P. J., Brasseur, R., Culbertson, M. R., Humphries, S. E., and Attie, A. D. (1994) *J. Biol. Chem.* **269,** 21670–21675
24. Benhizia, F., Ginsberg, H. N., Humphries, S. E., and Talmud, P. J. (2001) *Biochem. Biophys. Res. Commun.* **283,** 149–157
25. Byrne, C. D., Wareham, N. J., Mistry, P. K., Phillips, D. I. W., Martensz, N. D., Halsall, D., Talmud, P. J., Humphries, S. E., and Hales, C. N. (1996) *Atherosclerosis* **127,** 35–42
26. Prieur, X., Coste, H., and Rodriguez, J. C. (2003) *J. Biol. Chem.* **278,** 25468–25480
27. Khovidhunkit, W., Duchateau, P. N., Medzihradszky, K. F., Moser, A. H., Naya-Vigne, J., Shigenaga, J. K., Kane, J. P., Grunfeld, C., and Feingold, K. R. (2004) *Atherosclerosis* **176,** 37–44
28. Gonzalez-Conejero, R., Corral, J., Roldan, V., Martinez, C., Marin, F., Rivera, J., Iniesta, J. A., Lozano, M. L., Marco, P., and Vicente, V. (2002) *Blood* **100,** 2081–2086
29. Kozak, M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 8301–8305
30. Fullerton, S. M., Buchanan, A. V., Sonpar, V. A., Taylor, S. L., Smith, J. D., Carlson, C. S., Salomaa, V., Stengard, J. H., Boerwinkle, E., Clark, A. G., Nickerson, D. A., and Weiss, K. M. (2004) *Hum. Genet.* **115,** 36–56
31. Morgan, K., Scobie, G., and Kalsheker, N. A. (1993) *Hum. Mol. Genet.* **2,** 253–257
32. Li, W. W., Dammerman, M. M., Smith, J. D., Metzger, S., Breslow, J. L., and Leff, T. (1995) *J. Clin. Investig.* **96,** 2601–2605
33. Waterworth, D. M., Talmud, P. J., Luan, J., Flavell, D. M., Byrne, C. D., Humphries, S. E., and Wareham, N. J. (2003) *Biochim. Biophys. Acta* **1637,** 200–206
34. Nabika, T., Nasreen, S., Kobayashi, S., and Masuda, J. (2002) *Atherosclerosis* **165,** 201–204
35. Austin, M. A., Talmud, P. J., Farin, F. M., Nickerson, D. A., Edwards, K. L., Leonetti, D., McNeely, M. J., Viernes, H. M., Humphries, S. E., and Fujimoto, W. Y. (2004) *Biochim. Biophys. Acta* **1688,** 1–9
36. Baum, L., Tomlinson, B., and Thomas, G. N. (2003) *Clin. Genet.* **63,** 377–379