Southern blot analysis of the genomic DNA from 17 individuals showed perfect correspondence between a Bsp 1286I restriction fragment length polymorphism (RFLP) and the Ag(c/g) locus on human apolipoprotein (apo) B. The RFLP polymorphism is caused by a C→T transition at nucleotide 421 on the cDNA, resulting in a threonine-to-isoleucine conversion. Thus, three of the five Ag sites have now been tentatively located on the 4536 residue apo B peptide at amino acid residues 71: Ag(c/g), 591 (Ag(a1/d)), and 4154 Ag(t/z). (Arteriosclerosis 9:242–246, March/April 1989)

About 70% of total serum cholesterol is carried by low-density lipoproteins (LDL) in humans, and clinical studies have indicated a strong correlation between increased serum cholesterol levels and a higher risk of development of atherosclerosis. As the sole protein component on human LDL, apolipoprotein (apo) B plays a key role in lipid metabolism. For this reason, the genetic heterogeneity of apo B within the human population has become an important topic for investigation. The screening of humans for the existence of major restriction fragment length polymorphisms (RFLPs) has already provided abundant information about the genetic variation of apo B. Recently, Ma reported the existence of 14 RFLPs of apo B by using 23 restriction endonucleases and four large cDNA probes covering almost the entire coding sequence of apo B.

Antigenic polymorphisms of human apo B, on the other hand, were studied extensively in the 1960s and early 1970s. By studying the sera from patients who had received numerous blood transfusions, investigators identified a total of 10 epitopes that behave as five antithetical pairs: Ag(a1/d), Ag(c/g), Ag(h/l), Ag(x/y), and Ag(t/z). All of these Ag variants represent structural variations of apo B that can be recognized by antisera. Moreover, using the Ag markers, 15 major human apo B gene haplotypes have been identified, and frequency distributions have been measured in Australian aborigines, Bantu, Chinese, Indian, Senegalese, Swiss, and Tibetan populations. The existence of these 15 haplotypes, identifying 15 allelic major variants of the human apo B gene, provides a convenient ethnological, and perhaps historical, framework on which to place additional genetic variations, particularly those related to atherosclerosis.

Berg and his coworkers reported an association between the Ag(y) epitope and significantly higher levels of serum triglycerides and cholesterol in a study involving more than 1000 normal individuals. Tikkanen et al. reported a significant correlation between the Ag(c) epitope and higher cholesterol levels in 500 Finnish children. The Ag(c) locus is also recognized by a monoclonal antibody, Mb-19, which has been useful in studies relating the Ag system to apo B-100 and apo B-48. Certain apo B RFLPs have also been related to elevated serum lipids and to atherosclerosis.

We have been attempting to map the five Ag loci onto the apo B gene using RFLP analysis as well as molecular cloning techniques. We reported a perfect correlation for 17 individuals between the Ag(t/z) locus and an EcoRI RFLP near the 3' end of apo B gene, and this has been extended by 24 additional individuals. Recently, we identified a T→C substitution at nucleotide 1981 of apo B cDNA as being the polymorphic locus for Ag (a1/d). In this communication, we report another perfect association for 17 individuals (a complete linkage disequilibrium) between the Ag(c/g) locus and a new RFLP of apo B, detected by restriction enzyme Bsp 1286I. Combining the current and previous data, we may now tentatively locate three of these five Ag loci onto the genomic sequence of human apo B.

**Methods**

An ethnically diverse subpopulation of 17 unrelated normal individuals previously described was included in this study. DNA preparation, Southern analysis, and Ag phenotyping were also previously described. Briefly, genomic DNA from these 17 individuals was prepared from 30 ml of peripheral blood and was digested with various restriction enzymes under the conditions suggested by the manufacturers (New England Biolabs, Beverly, MA and Bethesda Research Laboratories, Gaithersburg, MD) but with at least a threefold excess of...
Table 1. Search for Restriction Site Length Polymorphisms Associated with Amino Acid
Sequence Changes in Apolipoprotein B

| Position  | Nucleotides alteration | Frequency | Restriction site predicted | No. alleles studied† | RFLP found | Correlation with Ag system |
|-----------|------------------------|-----------|----------------------------|----------------------|------------|---------------------------|
| 421       | C<->T                  | 2/6       | Bsp 1286I                  | 34                   | Yes        | Ag(c/g)                   |
| 1113      | C<->G                  | 1/8       | Hinf I                     | 8                    | No         | ‡                         |
| 1981      | C<->T                  | 2/6       | Alu I                      | 34                   | Yes        | Ag(a1/d)                  |
| 2506      | AG<->GA                | 1/6       | Awe NI                     | 8                    | No         | ‡                         |
| 2507      | 756                    |           |                            |                      |            |                           |
| 2883      | G<->C                  | 1/8       | Alu I                      | ND                   | §          |                           |
| 3453      | C<->G                  | 1/6       | Hae III                    | ND                   | §          |                           |
| 5727      | C<->G                  | 1/5       | Nco I                      | 8                    | No         | ‡                         |
| 6402      | G<->C                  | 1/5       | Psa I                      | ND                   | §          |                           |
| 6422      | C<->G                  | 1/5       | Alu I                      | 8                    | No         | ‡                         |
| 6790      | A<->T                  | 1/5       | (AATT)                     | ND                   | ‡          |                           |
| 7221      | A<->G                  | 1/5       | Hae III                    | 8                    | No         | ‡                         |
| 7748      | G<->A                  | 1/5       | Hae III                    | 8                    | No         | ‡                         |
| 8167      | A<->G                  | 1/5       | (AATT)                     | ND                   | ‡          |                           |
| 10 419    | G<->C                  | 1/5       | Msp I                      | 24                   | No         | §                         |
| 12 019    | 3937                   | 1/5       | (TGCA)                     | ND                   | ‡          |                           |
| 12 459    | A<->C                  | 1/5       | Nla III                    | ND                   | ‡          |                           |
| 12 511    | 4101                   | 1/6       | Taq I                      | 8                    | No         | ‡                         |
| 12 528    | 4106                   | 1/6       | Bbv I                      | ND                   | ‡          |                           |
| 12 669    | 4154                   | 2/6       | EcoRI                      | 34                   | Yes        | Ag(t/z)                   |

*As listed in Ludwig et al.†
†Number of persons studied > 2. ‡Not Ag(x/y) or Ag(c/g). §Polymorphic fragments not distinguishable. |No enzyme available for this palindrome. †This enzyme is very expensive.
RFLP = restriction fragment length polymorphisms, ND = not done.

enzyme. After digestion, 10 µg of DNA from each person was separated through gel electrophoresis by using 1% to 2% agarose depending on the size of the restriction fragments. This was blotted onto nylon filters (NyTran, 0.2µm, Schleicher & Schuell, Keene, NH). The apo B cDNA probe was doubly radiolabeled with alpha-32P-dATP and alpha-32P-dCTP to specific activities of greater than 3 x 106 cpm/µg. Hybridization and autoradiography were performed as described.17 The experimental protocol was approved by the Human Subjects Protection Committee of the University of California, Los Angeles.

Results

Recently, Ludwig et al.19 sequenced the genomic DNA of human apo B and compared the apo B gene sequence with four complete and three partial cDNA sequences of apo B previously reported by other authors.20-26 Ludwig and coworkers found a total of 60 nucleotide substitutions and 39 amino acid substitutions. Although some of these differences might be due to errors in DNA sequencing, others must be real. Those substitutions involving amino acid changes were of special interest to us since some of them might reflect the Ag polymorphisms. We then examined all of these apo B sequences, and found a total of 16 base substitutions involving restriction sites as shown in Table 1. These restriction sites include the site at nucleotide 1981 for the Ag(a1/d) polymorphism involving an Alu I RFLP and the site at nucleotide 12 669 for the Ag(t/z) polymorphism involving an EcoRI RFLP, as previously reported.17,18

Our attention next became focused on a search for RFLPs that might correspond to one of the other Ag pairs. Based on the availability of the restriction enzymes, as well as the size of the predicted restriction fragments, we tested 11 of the 16 possible RFLPs with corresponding restriction enzymes (Table 1) and appropriate probes, using DNA from four individuals who were informative for the Ag(c/g) and the Ag(x/y) loci (i.e., a group including both homozygous and heterozygous individuals for these two loci). One of these RFLPs, detected by the Bsp 1286I restriction enzyme, matched the Ag(c/g) polymorphism for these four individuals. None of these 11 predicted RFLPs matched the pattern for Ag(x/y).

The Bsp 1286I RFLP was detected using a 5'-end probe, EL-4, composed of cDNA nucleotides 123 to 1235,27 and the restriction sites involved in generating the fragments are shown in Figure 1. The Southern analyses for five individuals are shown in Figure 2. The base change responsible for this RFLP is located at nucleotide 412, and involves a C-to-T conversion resulting in a threonine-to-isoleucine substitution at amino acid residue 71. The Ag(c) epitope is associated with the threonine and the presence of the restriction site, yielding two fragments of 141 and 114 base pairs (bp). The Ag(c) epitope is associated with the absence of the restriction site, yielding a 255-bp band in the Southern analysis. A more detailed study of 17 individuals who had been previ-
Figure 1. Restriction map of Bsp 1286I restriction site length polymorphism. Three GNGCN'C sequences indicate the restriction sites of Bsp 1286I, where N may be any nucleotide, and ' marks the cutting site. The shaded box indicates exon 4 of the apoprotein B gene, and the thin line adjacent to the box indicates the intron sequence. The cDNA probe EL-4 extends from cDNA nucleotides 123 to 1235; thus, it begins in the 5' noncoding region and completely covers the first eight exons and most of the ninth. The polymorphism at nucleotide 421 is a C (presence of the cutting site) to T conversion.

Discussion

It is known that Ag(c/g) lies within thrombin cleavage fragment T4,\(^2^9\) which extends to amino acid residue 1297.\(^2^4\) In this region, Ludwig et al.\(^1^8\) tabulated eight additional polymorphisms, which result in amino acid changes among the six published sequences that they examined. From the data presented in Table 1, the sites at amino acid residues 302, 591, and 766 may be eliminated as possible Ag(c/g) sites. The possibility remains that one of the other five sites results in the structural change in apo B actually recognized by the human allotypic antibodies defining Ag(c/g), but if so, it clearly must be tightly linked to the polymorphism at amino acid residue 71. We have recently learned that Stephen Young of the Gladstone Foundation (personal communication) has independently found the same perfect correlation between this polymorphism and the Ag(c/g) locus in a population of
Table 2. Association between Ag System Loci and Bap 12861 Restriction Site Length Polymorphisms

| Donor | a1 | d | g | c | h | i | x | y | l | z | B1 | B2 |
|-------|----|---|---|---|---|---|---|---|---|---|----|----|
| 01    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 02    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 03    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 04    | -  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 05    | -  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 06    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 07    | -  | - | - | - | - | - | - | - | - | -  | -   | -   |
| 08    | +  | - | - | - | - | - | - | - | - | -  | -   | -   |
| 09    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 10    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 11    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 12    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 13    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 14    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 15    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 16    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |
| 17    | +  | + | + | + | + | + | + | + | + | +  | +  | +  |

Freq*: 0.50 0.50 0.85 0.35 0.06 0.94 0.21 0.79 0.74 0.26 0.65 0.35
Freq†: 0.50 0.50 0.70 0.30 0.06 0.94 0.25 0.75 0.80 0.20 NA NA

*For the Ag loci and the RFLP, the frequencies listed are those determined from the 17 individuals in this study. Given the small sample size, these are not significantly different from those reported for the Swiss population listed below.
†For the Ag loci, the frequencies listed are those determined from a Swiss population of 382 individuals. The frequencies for the Bap 12861 RFLP in a large population are not available at this time.
‡Not tested but assumed positive because of the rarity of the homozygote for Ag(h) allele.

Table 3. Distribution of Ag Phenotypes among Bap 12861 Genotypes

| DNA polymorphism | Genotype | Ag(c/g) | Ag(d/g) | Ag(g/g) |
|------------------|----------|---------|---------|---------|
| Bap 12861        | B1B1     | 1 0 0   | 0       | 0       |
|                  | B1B2     | 0 10 0  | 0       | 0       |
|                  | B2B2     | 0 0 0   | 6       |

\(\Delta=1.00 \text{ with } 95\% \text{ confidence limits } 0.88 < \Delta < 1.00 \text{ (A. Chekravarti, personal communication).}\)

62 unrelated persons. Together with our study, this means that 79 individuals or 158 alleles are perfectly matched. If \(p\) is the average probability that a single allele taken at random will match, then the lower 95% confidence limit on \(p_{\text{pooled}}\) can be calculated as \(p_{\text{pooled}}>0.98\). Clearly, this RFLP will be useful for Ag(c/g) haplotyping regardless of whether or not it coincides with the nucleotide change responsible for the immunologic polymorphism.

A perfect match of 17 out of 17 individuals was previously found between the EcoR I RFLP and the Ag(t/z) polymorphism, and this perfect match has now been extended by 24 additional individuals increasing the total to 41. This RFLP has been located in the coding sequence for apo B at nucleotide 12,669 involving an A-to-T nonsynonymous substitution converting a lys to glu. The nature of the RFLP and the perfect match suggested that the EcoR I RFLP might be the actual cause of the Ag(t/z) polymorphism, or at least that the RFLP was closely linked to the Ag(t/z) locus.

We previously conducted a detailed study of the molecular basis of the Ag(a1/d) polymorphism using molecular cloning techniques and RFLP analysis. An Alu I RFLP was found to be perfectly correlated with the Ag(a1/d) site. This is a C-to-T substitution at nucleotide 1981, which changed the amino acid sequence from an alanine in the Ag(d/d) individuals to a valine in the Ag(a1/a1) individuals.

Combining the previous Ag(t/z) and Ag(a1/d) data, we are now able to tentatively locate three out of the five Ag loci along the apo B gene (Figure 3).
