The human apolipoprotein A-II gene: complete nucleic acid sequence and genomic organization

Karl J. Lackner*, Simon W. Law and H. Bryan Brewer, Jr.

Molecular Disease Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20205, USA

Received 25 February 1985; Revised 9 May 1985; Accepted 10 May 1985

ABSTRACT

The gene for human apolipoprotein (apo) A-II has been isolated from a human genomic DNA library. The cloned fragment was approximately 14 kilobase-pair (kb) long, and extended about 9.0 kb upstream as well as 3.5 kb downstream from the apoA-II gene, which was contained within a 3.1 kb HindIII fragment of human DNA. The complete nucleic acid sequence of the apoA-II gene has been determined, establishing that the apoA-II gene is interrupted by three intervening sequences of 182, 293, and 395 bp. The second intron is of particular interest, because it contains a 33 bp sequence of alternating G and T residues very close to the 3' splice site which has the potential to form a left handed Z-helix structure in vivo. A restriction fragment length polymorphism 3' from the apoA-II gene has been detected which may serve as a marker for the long arm of chromosome 1 in linkage analyses.

INTRODUCTION

Apolipoprotein (apo) A-II is a major protein constituent of human high density lipoproteins (HDL) (1). In plasma apoA-II is found predominantly as a 17,400 dalton protein composed of two identical 77 amino acid polypeptides covalently linked by a single disulfide bond between the cysteines at position 6 (2). ApoA-II is synthesized by the liver and intestine as a precursor protein (3). In vitro translation studies (3) and nucleic acid sequence analysis of apoA-II cDNA clones (4-7) revealed that apoA-II mRNA encodes a 100 amino acid long preproprotein which includes an 18 amino acid prepeptide, a 5 amino acid propeptide, and the 77 amino acid mature protein. The prepeptide is cleaved co-translationally, while the propeptide is post-translationally removed. This latter cleavage has been reported to occur extracellularly in the human hepatoma cell line HepG2 (3,8). Since the propeptide of apoA-II terminates with two positively charged arginine residues, a common feature of most intracellularly cleaved proproteins (9), the extracellular cleavage of proapoA-II may be unique, and of physiological significance in apoA-II metabolism.

Epidemiologic studies have indicated that there is an inverse correlation

between plasma HDL-cholesterol levels and the risk for development of premature cardiovascular disease (10,11). In addition, the two major apolipoproteins of HDL, apoA-I and apoA-II, have been reported to be independent risk factors which are inversely correlated with premature atherosclerosis (12). The plasma levels of HDL and apoA-I have also been proposed to be useful as independent predictors of the severity of coronary artery disease in patients with angina (13,14). These results suggest that analysis of the molecular mechanisms involved in the regulation of the plasma levels of HDL, apoA-I, and apoA-II will provide a better understanding of the genetic factors which predispose to the development of premature heart disease.

An improved knowledge of the genetics or inheritance of HDL, apoA-I, and apoA-II may also provide additional markers for individuals at an increased risk for premature cardiovascular disease. We have previously reported the cloning of cDNAs for human apoA-I (15,16), and apoA-II (5) from a human liver cDNA library, and the use of these cDNAs to localize the respective genes to single chromosomes (17,18). The apoA-I gene was found to be localized to chromosome 11 close to the centromere, while the apoA-II gene was located on chromosome 1 in the region 1p21--> 1qter. These results preclude genetic linkage between the two major apolipoproteins in human plasma HDL. In this report, we present the complete nucleic acid sequence and genomic organization of the apoA-II gene.

MATERIALS AND METHODS

Isolation of genomic DNA from white blood cells.

White cell DNA was prepared as previously described (19,20). Briefly, 90 ml of 0.32M sucrose, 5mM MgCl₂, 10 mM Tris-HCl (pH 7.5), and 1% Triton X-100 were added to 10 ml of blood anticoagulated with EDTA (1mg/dl) at 4°C. The sample was centrifuged at 1000xg for 15 minutes and nuclear pellets were resuspended in 4.5 ml of 24mM EDTA, 75 mM NaCl (pH 8.0), and 0.5ml each of 5% sodium dodecylsulfate (NaDodSO₄) and proteinase K (2mg/ml, Boehringer-Mannheim) added. The solution was then incubated at 37°C for 12-16 hrs, followed by sequential extractions with an equal volume of solutions containing phenol, phenol/chloroform/isoamylalcohol (50:24:1), and chloroform/isoamylalcohol (24:1). Finally, following extraction, sodium acetate was added to a final concentration of 0.3M and the DNA precipitated with ethanol.

Screening of a genomic library for apoA-II clones.

A human placental genomic DNA phage library was a generous gift of Dr. Philip Leder. E.coli C-600 was infected with the Charon 28 recombinant phage,

and plated onto 150 mm round agar plates, at approximately 15-20,000 plaques per plate. Phage plaques were lifted with nitrocellulose paper (Schleicher & Schuell). The DNA was denatured with 0.5 N NaOH containing 1.5 M NaCl for 2x2 min, neutralized with 0.5 M TrisHCl and 1.5 M NaCl (pH 7.4) for 2x5 min, rinsed briefly with 2x SSC, (SSC: 150 mM NaCl, 15 mM Na Citrate, pH 7.1) and baked for 2 hrs. at 80°C. The apoA-II specific cDNA probe used to screen the library was a 300 bp fragment. It was generated by cleavage of the plasmid pMDB2049 (5) with PstI and XbaI. The probe was labelled by nick translation with (α -³²P)dCTP (New England Nuclear) to high specific activity ($3\text{--}5 \times 10^8$ dpm/ug). The nitrocellulose paper was prewashed for 1 hour in 50 mM Tris-HCl, 1 M NaCl, 1 mM EDTA, and 0.1% NaDodSO₄ (pH 8.0) at 68°C, prehybridized for 4 hrs in 6x SSC, 5x Denhardt's, 0.1% NaDodSO₄, and 100 ug/ml sonicated salmon sperm DNA at 68°C, and hybridized in the latter solution with 10⁷ dpm/filter of the nick translated apoA-II probe. After hybridization the filters were washed in 1xSSC, and 0.1% NaDodSO₄ at 68°C for 2x30 min. Positive clones were selected and rescreened as outlined above at lower density until single pure clones could be isolated.

Large scale preparation of phage DNA.

Large scale preparations of phage DNA were performed as described by Maniatis et al. (21) with minor modifications. 0.5 l of L-Broth containing 10 mM MgSO₄ was inoculated with approximately 10¹⁰ E. coli C-600 previously infected with 5 X 10⁸ plaque forming units of a single clone. After 14-16 hours incubation at 37°C, 12 ml of chloroform was added to the culture and the incubation extended for another 30 minutes. After addition of NaCl to 1M, the culture was incubated at 0°C for 1 hour, and the bacterial debris removed by centrifugation at 11,000xg for 10 minutes at 4°C. The phage were then precipitated by addition of polyethylene glycol (MW 8,000; Sigma) to 10% (w/v). After centrifugation (11,000xg, 10 min, 4°C) the pellets were resuspended in 100 mM NaCl, 10mM MgSO₄, 50 mM Tris-HCl, and 0.01% gelatine (pH 7.5) (SM buffer) an equal volume of chloroform added, and the remaining debris removed by a 15 minute centrifugation at 1600xg (4°C). To each ml of the aqueous supernatant 0.75 g of CsCl was added and the phage purified by ultracentrifugation in a Beckman 40.3 rotor at 31,000 rpm for 24 hours at 4°C. The band containing the phage particles was collected and dialyzed against 10mM NaCl, 50mM TrisHCl, and 10 mM MgCl₂ (pH 8.0). Phage proteins were digested with proteinase K (50 ug/ml) at 65°C for 1 hour and removed by phenol extraction. The purified phage DNA was recovered by ethanol precipitation.

Subcloning into pBR322.

A 3.1 kb HindIII fragment of the insert genomic DNA isolated from λ A-II-274 was purified by agarose gel electrophoresis and inserted into the HindIII site of pBR322. Recombinant clones were selected by testing for antibiotic resistance and hybridization to the apoA-II cDNA probe. Plasmid DNA was isolated as described previously (5).

Restriction enzyme digests.

Restriction endonucleases were purchased from New England Biolabs or Bethesda Research Laboratories. Reactions were performed under the conditions specified by the suppliers. Cleavage of genomic DNA was performed with 5 units of enzyme per μ g of DNA for 4 hours.

Southern blot analysis.

Transfer of DNA to nitrocellulose filter paper was performed as previously reported (22,23). Hybridization to the cDNA probes was performed as described above for the library screening except that the prewash was omitted.

DNA sequencing. DNA fragments were end labelled as previously reported (24,25), and sequenced by the chemical cleavage method of Maxam and Gilbert (25).

RESULTS

High molecular weight DNA was isolated from 18 individuals. Restriction endonuclease digests were analyzed by the Southern blot technique using as hybridization probe a 300 bp long PstI/XbaI fragment which covered the sequence from the codon for amino acid -10 into the 3' untranslated end of our previously described apoA-II cDNA clone pMDB2049. We detected one restriction fragment length polymorphism (RFLP) with MspI (Fig. 1). This enzyme generates either a 3.2 kb or a 3.8 kb fragment. The polymorphic site has been determined to be located 3' from the apoA-II gene by analyzing MspI/SstI double digests. In our sample we identified 11 homozygotes for the small fragment, 2 homozygotes for the large fragment, and 5 heterozygotes. The small sample size allows only for an approximate determination of the allelic frequency, which is estimated to be approximately 0.75 for the 3.2 kb fragment and 0.25 for the 3.8 kb fragment, respectively. This fairly common RFLP should be useful as a marker for the long arm of chromosome 1 to which the apoA-II gene has been assigned (18) in genetic linkage analysis. The results of the restriction digests enabled us to construct a restriction endonuclease cleavage map for the apoA-II gene as shown in Fig. 2A. This map was initially based on the assumption that the recognition sites for PstI, which is in the coding region

Figure 1: Southern blot of human DNA cleaved with MspI. There are five homozygotes for the 3.2 kb fragment of the apoA-II gene, two heterozygotes, and one homozygote for the 3.8 kb fragment. Molecular weight markers are included on the vertical axis.

for the leader sequence of apoA-II cDNA, and SstI, which is in the 3' untranslated region, were also present in the equivalent regions of the genomic DNA. This assumption was later verified by the sequence analysis of the genomic clone. The restriction map was generated using these two cleavage sites as reference points. The positions of the PstI and SstI recognition sites within a 3.1 kb HindIII fragment indicated that the entire apoA-II gene was present in this fragment.

Figure 2: Restriction Map of the ApoA-II Gene. A. Restriction sites in the structural and flanking region of the apoA-II gene. The polymorphic MspI site is indicated by an asterisk. B. Restriction sites within the fragment sequenced. Only enzymes used to generate fragments for end labelling are presented. Sequences which are part of the spliced mRNA are shown as wide bars. C. Strategy for sequencing the apoA-II gene. The arrows illustrate the direction and length of the fragments sequenced. The upper two lines represent 3' labelled fragments. All other fragments were 5' labelled.

The apoA-II genomic clones were identified in a human genomic λ -phage library by screening approximately 1.5×10^6 phage plaques with the apoA-II cDNA probe. Two positive clones were identified. One clone, designated λ A-II-274, was grown in large scale and analyzed following restriction endonuclease digestion by Southern blot. This clone contained a 14 kb fragment of human genomic DNA inserted into the BamHI cloning site of Charon 28, and analysis with restriction enzymes detected the 6.9 kb SstI and the 3.1 kb HindIII fragments present in genomic DNA indicating that there were no major rearrangements in λ A-II-274. Further analysis revealed that the insertion extended about 9.0 kb upstream and 3.5 kb downstream from the apoA-II gene. The 3.1 kb HindIII fragment containing the apoA-II gene was purified by agarose gel electrophoresis, and subcloned into the HindIII site of pBR322. One recombinant which contained a single copy of the HindIII fragment was used for restriction enzyme as well as sequence analysis.

```

GCCTCCACTAAGAGCTGTACCCCTGCCTCTCACCCATCACCATGAGTCTTCATGTGCTTGTCTCTCTCTCCCC 75
CATTTCTCCAACCTGTTTATCCTCACATAATCCCGCCCCACTGGGCCCATCCATAGTCCCTGTCACTTGACAGGG 160
GGTGGGTAACAGACAGGTATATATGCCCCCTTCTCTCCAGCCAAGGCAGGCACAGACACCAAGGACAGAGACGCT 225
GGCTAGGTAAAGATAAGGAGGTAAAGATAAGGAGGCAAGATGTGTGAGCAGCATCCAAAGAGGCTTGGGCTTCAGTT 300
GTGGAGAGGGAGAGAGCTAGGTTGGAATGGGCAGCAGGTAGGGAGATCCCTGGGGAGGAGCTGAAGCCCATTTGG 375
CTTCAGTGTCCCCCAAACCCCAACCCCTCTTCTTAGGCGGCCCTCCCACTGTTACCAACATGAAGCTGCTCG 460
MetLysLeuLeuA
CAGCAACTGTGCTACTCTCACCATCTGCAGCCTTGAAGGTGGGTGTGATATGGAGAGGGGCCAAAGAGGAGAA 525
LeuAlaThrValLeuLeuLeuThrIleCysSerLeuGluG
TGTGCTGGTGGTTATAGCTACCTATCTGCCTGTGCTCTTATATCCAGGTCTGGCAACCAAGGGCTCAAAAGGAA 600
GATCATTCCCTCTCTAAAGCCAGAAAGCCCATCCAAAGATCTGGCTGGATACCTTTTTGGGGAGGGGGCAGAGA 675
GCTGGGGTGGGTACTGGCAGAGGTATGGCAGGAGGCCAGGATTCACTGTGTGGACCCAGCTGAAAAGAGTGTGT 760
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 825
IyAlaLeuValArgArgGlnAlaLysGluProCysValGluSe
CCTGGTTTCTCAGTACTTCCAGACCGTGACTGACTATGGCAAGGACCTGATGGAGAAGGTCAAGAGCCCAAGCT 900
rLeuValSerGlnTyrPheGlnThrValThrAspTyrGlyLysAspLeuMetGluLysValLysSerProGluLe
TCAGGCCGAGGCCAAGTAAGTCTCAGGGCAAGGGGTTCAGGGGCTGTGGAACGTGTGGAGAGAAAGAAGGGAAGAT 975
uGlnAlaGluAlaLy
GAGAGGTCCACAGAAGTCTGAACCCAGGGGTGGGGATTAGGGCAGATTAGGCTTAAATTGCAGAGAAAAAGTAT 1060
TTCATCACCCAAAGATCCACACGCTCTTATAGATAGAGAGGAACAGCAAGAACTGGGCCTTGAATTTAGTCTCT 1125
AGAGTCTGTCCCTCTACCTAGCAAAGGTCTTGACTCTATTCTACCTAGGGGCTTTGCCATGCGATGGACCAGGC 1200
ACTAGAGTTTGGGACCTGAGTCAGTCTGCTCTGACCTCCACCACCACCAAGGCCCTGCCAGTGCCTAGGGTC 1275
CCTCAGATTAACTCTAATCCCTCACCTATCCAGGTCTTACTTTGAAAAGTCAAAGGAGCAGCTGACACCCCTG 1360
sSerTyrPheGluLysSerLysGluGlnLeuThrProLeu
ATCAAGAAGGCTGGAACGGAAGTGGTTAACTTCTTGAGCTATTTCTGTGGAACCTTGAACACAGCCTGCCACCCAG 1425
IleLysLysAlaGlyThrGluLeuValAsnPheLeuSerTyrPheValGluLeuGlyThrGlnProAlaThrGln
TGAAGTGTCCAGACCAATTGCTCTTCCAAACCCAGCTGGCCTCTAGAACACCCACTGGCCAGTCTCTAGAGCTCCTGT 1600
***
CCCTACCCACTCTTTGCTACATATAATGCTGAATGAATCCAGCTCTGAGCCTGGTATGTTTGGGGGACTGGGAAA 1675
AGTAGGGGAGTAAGGGAGGAGAGAGGAAAGGAAAGGAAATCTGCTTCTAGAAGGAGAGAGGTTTGAGTGTGG 1680
AGGGGTGAAGAAAGGATTGAAGACACAACATGATGAAATGACAGGATGAGGGTGCCATGATTCTCCAAACCCAGA 1725
GTCTCTACAGCTTGGGCACGACTCTGCAG

```

Figure 3: Complete Nucleotide Sequence of the ApoA-II Gene. Sequences present in the spliced mRNA are underlined, and the encoded amino acids are indicated underneath the respective nucleotide triplets. *** indicates the termination codon. Three regions of interest have been boxed. Nucleotides 119-127 which are very similar to the consensus sequence for the CAAT-box, nucleotides 169-174 which represent a typical Goldberg-Hogness box, and the polyadenylation signal at nucleotides 1521-1526. The complementary sequences between nucleotides 133 and 150 are underlined. The open triangle indicates the cleavage site of the leader sequence. The solid triangle indicates the cleavage site between proapoA-II and mature apoA-II.

A restriction map of the region of the apoA-II gene within the 3.1 kb HindIII fragment is presented in Figure 2B, and Figure 2C illustrates the strategy employed in the sequence analysis of the apoA-II gene. Fragments were 3' end labelled if they contained protruding 3' ends, or 5' end labelled when they had blunt or protruding 5' ends. The nucleotide sequence of each

Figure 4: Schematic Representation of the Transcription and Translation of the ApoA-II Gene. The upper section of this figure illustrates the apoA-II gene. Exons are shown as open bars, and introns are represented as solid bars. * indicates the position of the CAAT-box, ** indicates the TATA-box, *** indicates the polyadenylation signal. The GT repeats at the end of the second intron are hatched. The middle section of the figure contains the spliced mRNA including the splice sites. The lower part illustrates the translated preproapoA-II. The presequence from amino acid -23 to -6 is hatched, the propeptide from -5 to -1 is open, and the mature plasma protein from 1 to 77 is stippled.

fragment was obtained by the Maxam and Gilbert procedure (25).

The complete nucleotide sequence of the apoA-II gene is shown in Figure 3. A schematic overview summarizing the structure of the apoA-II gene, its mRNA, and the synthesized protein is presented in Figure 4. A typical TATA box structure is located from 29 to 24 bp upstream of the 5' end of the mRNA. Seventy-nine bp upstream of the mRNA start site is a nucleotide sequence which matches the consensus sequence for the CAAT-box (26) in 7 out of 9 positions.

The apoA-II gene is interrupted by three intervening sequences. The first one is 182 bp long and is located between nucleotides 34 and 35 of the 5' untranslated end of the apoA-II mRNA. The second intron, 293 bp in length, is located between the first and second nucleotides of the codon for amino acid -6 of the signal peptide, and separates the prepeptide from the rest of the apoA-II protein sequence. This intron contains a 33 bases long sequence of GT repeats terminating only six bp from the 3' splice site. The third intervening sequence is the longest with 395 bp, and is located between the second and third base of the codon for amino acid 39 of the mature plasma apoA-II apolipoprotein.

DISCUSSION

In the present report we have identified, isolated, and characterized the human gene for apoA-II. The gene is contained in a 3.1 kb HindIII fragment of human chromosome 1 (18). It is interrupted by three intervening sequences.

Previous studies in our laboratory have shown that the 5' untranslated end of apoA-II mRNA is 58 bases long (5). This is at variance with the results of Knott et al (7) who reported the 5' untranslated end to be 61 bases long. Based on our data nucleotide 198 in the genomic sequence is the mRNA start site. This nucleotide is an adenosine which is the most commonly reported nucleotide at the mRNA start site (26). Twenty nine basepairs upstream from this nucleotide is a typical Goldberg-Hogness box (26) at nucleotides 169-174. This result is consistent with nucleotide 198 being the mRNA start site, since the TATA box begins in more than 80% of the known human genes from 33 to 29 bp upstream from the mRNA start site (24).

Another putative promoter region, the CAAT-box, has been detected in several eucaryotic genes 70-80 bp upstream from the mRNA start site (26,27). The apoA-II gene contains a nucleotide sequence from 119 to 127 which is very similar to the consensus sequence for the CAAT-box (5' GG^C_TCAATCT 3' vs. 5' GGGCCATCC 3'). The only differences are the two cytosines at positions 123 and 127. Recent analysis of the influence of single base substitutions in this region in the rabbit β -globin gene revealed that a T to C substitution in the last position of the CAAT-box did not affect the relative level of gene expression, whereas several other substitutions in this sequence profoundly decreased gene expression (28). An A to C transversion in the fifth position of the CAAT-box, as is present in the sequence of the apoA-II gene, was not analyzed by these investigators (28).

Another interesting feature of this region in the apoA-II gene is the sequence between nucleotides 133-150 (Fig.3). Nucleotides 133-140 are complementary to nucleotides 150-143, and thus have the potential to form secondary structure. Whether the location of this structure between the CAAT-box and the TATA-box has any relevance to the control of the expression of the apoA-II gene remains to be investigated.

The 5' untranslated sequence of apoA-II mRNA as determined from our genomic clone is in complete agreement with our previously published sequence determined by primer extension (5). However our sequence is different in several positions from the 5' untranslated end of the apoA-II mRNA as reported by Knott et al. (7). We have sequenced three fragments covering this region, and found a single unambiguous sequence for the first exon. At present we do not

know the reason for these differences. It is interesting, however, to note that all these differences occur in the first exon. This could thus represent true genetic heterogeneity in the first exon of the gene for apoA-II.

The second intervening sequence of the apoA-II gene contains a 33 bp long sequence of GT repeats only 6 bp upstream from the 3' splice site. Long stretches of alternating purine and pyrimidine residues have been shown to have the potential to form left handed helix or Z-DNA structure (29). Even though GT or AC repeats do not form Z-DNA helices as readily as GC repeats, it is now well established that they also have the potential to form left handed helices. GT repeats are fairly abundant in the mammalian genome (30,32), and it has been speculated that these sequences may be associated with an increased frequency of recombination (33). This intervening sequence separates the signal- or prepeptide from the proform of apoA-II (Fig. 4) which has been proposed to be the secretory form of apoA-II (3). Thus it is conceivable that the codons for the prepeptide and the secreted protein were linked by a recombination process during evolution, which had been facilitated by the Z-DNA structure present at the end of the second intron. Furthermore this sequence may be functioning in the regulation of transcription, by changing its conformation (29). To our knowledge a potential Z-DNA structure has not been reported in such close proximity to a mRNA splicing site. The effect of this GT-tract on mRNA splicing is not known. Most likely it will not interfere with the normal splicing process. However, the existence of such a uniform and simple sequence between the 3' splice site and the potential branch point for the lariat formation supports the concept that only narrow regions in the intron are involved in the splicing process.

It has been postulated based on comparison of the primary structure of the plasma apolipoproteins that apoA-II, apoA-I, as well as apoC-I and apoC-III were derived from a common ancestor (34). Studies on protein and DNA sequences in rat apoA-I, apoA-IV, and apoE supported this concept (35). The structure of the apoA-II gene as presented here is quite similar to that of apoA-I (Fig. 5) (4,36). Both apolipoproteins have a short first exon of approximately 30 bp in the 5' untranslated end of the mRNA. The next exon codes for the majority of the signal peptide. The third exon contains the propeptides and the first 38 or 42 amino acids of apoA-II and apoA-I respectively. The fourth exon contains the remainder of the mRNA (Fig. 5). It is interesting to note, that the genomic organization of apoC-III also follows this pattern (37). The similarity of the genomic structure of apoA-II, apoA-I and apoC-III supports the concept that the genes were derived from a common ancestral gene.

Figure 5: Comparison of the ApoA-I and ApoA-II mRNAs. The spliced mRNAs are represented with the triangles indicating the sites of the intervening sequences. The narrow bars represent the noncoding sequences. The hatched portions of the wide bars represent the coding sequence of the prepeptides, the open portions that of the propeptides, and the stippled portion the mature plasma protein.

The availability of genomic and cDNA clones for the two major apolipoproteins in HDL will now permit the investigation of the molecular mechanisms involved in the regulation of apoA-I and apoA-II biosynthesis. Additionally, the RFLP using *Msp*I will make linkage analysis between the apoA-II gene and dyslipoproteinemias possible. This approach has recently led to the linkage of an apoA-I RFLP and hypertriglyceridemia (38). Thus, it may now be possible to define additional genetic loci involved in lipoprotein expression and metabolism.

ACKNOWLEDGEMENTS

K.J. Lackner is recipient of a research training fellowship from the Deutsche Forschungsgemeinschaft La 499/1-1. The authors wish to thank D.Mok and E. Schweins for technical assistance, and Dr. A. Dugaiczky for helpful discussion on the screening of the phage library.

*Address all correspondence to: Dr. Karl J. Lackner, Bldg. 10, Room 7N117, National Institutes of Health, Bethesda, MD 20205, USA

REFERENCES

1. Kostner, G.M., Patsch, D.R., Sailer, S., Braunsteiner, H., and Holasek, A. (1974) *Eur. J. Biochem.* 45, 611-621.
2. Brewer, H.B. Jr., Lux, S.E., Ronan, R., and John, K.M. (1972). *Proc. Natl. Acad. Sci. USA* 69, 1304-1308.
3. Gordon, J.I., Budelier, K.A., Sims, H.F., Edelstein, C., Scanu, A.M., and Strauss, A.W. (1983). *J. Biol. Chem.* 258, 14054-14059.
4. Sharpe, C.R., Sidoli, A., Shelley, C.S., Lucero, M.A., Shoulders, C.C., and Baralle, F.E. (1984) *Nucl. Acids Res.* 12, 3917-3932.
5. Lackner, K.J., Law, S.W., and Brewer, H.B. Jr., (1984) *FEBS Lett.* 175, 159-164.
6. Moore, M.N., Kao, F.T., Tsao, Y.K., and Chan, L. (1984) *Biochem. Biophys. Res. Comm.* 123, 1-7.

7. Knott, T.J., Priestley, L.M., Urdea, M., and Scott, J. (1984) *Biochem. Biophys. Res. Comm.* 120, 734-740.
8. Lackner, K.J., Edge, S.B., Gregg, R.E., Hoeg, J.M., and Brewer, H.B. Jr., (1985) *J. Biol. Chem.* 260, 703-706.
9. Docherty, K., and Steiner, D.F. (1982) *Annu. Rev. Physiol.* 44, 625-638.
10. Castelli, W.P., Doyle, J.T., Gordon, T., Hames, C.G., Hjortland, M.C., Hulley, S.B., Kagan, A., and Zukel, W.J. (1977) *Circulation* 55, 767-772.
11. Avogaro, P., Bittolo Bon, G., Cazzolato, G., Quinci, G.B., and Belussi, F. (1978) *Artery* 4, 385-394.
12. Fager, G., Wiklund, O., Olofsson, S.-O., Wilhelmssen, L., and Bondjers, G. (1981) *Arteriosclerosis* 1, 273-279.
13. Jenkins, P.J., Harper, R.W., and Nestel, P.J. (1980) *Brit. Med. J.* II. 187-192.
14. Maciejko, J.J., Holmes, D.H., Kottke, B.A., Zinsmeister, A.R., Dinh, D.M., and Mao, S.J.T. (1983) *New Engl. J. Med.* 309, 385-389.
15. Law, S.W., Gray, G., and H.B. Brewer, Jr., (1983) *Biochem. Biophys. Res. Comm.* 112, 257-264.
16. Law, S.W., and Brewer, H.B. Jr. (1984) *Proc. Natl. Acad. Sci. USA* 81, 66-70.
17. Law, S.W., Gray, G., Brewer, H.B. Jr., Sakaguchi, A.Y., and Naylor, S.L. (1984) *Biochem. Biophys. Res. Comm.* 122, 877-883.
18. Lackner, K.J., Law, S.W., Brewer, H.B. Jr., Sakaguchi, A.Y., and Naylor, S.L. (1984) *Biochem. Biophys. Res. Comm.* 122, 877-883.
19. Kunkel, L.M., Smith, K.D., Boyer, S.H., Borgaonkar, D.S., Wachtel, S.S., Miller, O.J., Breg, W.R., Jones, H.W., Jr., and Rary, J.M. (1977) *Proc. Natl. Acad. Sci. USA* 74, 1245-1249.
20. Bell, G.I., Karam, J.H., and Rutter, W.J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 5759-5763.
21. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning, a Laboratory Handbook*, Cold Spring Harbor Laboratory, New York.
22. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.
23. Meinkoth, J., and Wahl, G. (1984) *Anal. Biochem.* 138, 267-284.
24. Yousaf, S.I., Carroll, A.R., and Clarke, B.E. (1984) *Gene* 27, 309-313.
25. Maxam, A.M., and Gilbert, W. (1980) *Meth. in Enzymol.* 65, 499-560.
26. Breathnach, R., and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349-383.
27. Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) *Nucleic Acids Res.* 8, 127-142.
28. Dierks, P., Van Doyen, A., Cochran, M.D., Dobkin, C., Reiser, J., and Weissman, C. (1983) *Cell* 32, 695-706.
29. Rich, A., Nordheim, A., and Wang, A.H.-J. (1984) *Ann. Rev. Biochem.* 53, 791-846.
30. Nishioka, Y., and Leder, P. (1980) *J. Biol. Chem.* 255, 3691-3694.
31. Hamada, H., Petrino, M.G., and Kakunaga, T. (1982) *Proc. Natl. Acad. Sci. USA* 79, 6465-6469.
32. Slightom, J.L., Blechl, A.E., and Smithies, O. (1980) *Cell* 21, 627-638.
33. Kilpatrick, M.W., Klysik, J., Singleton, C.K., Zarling, D.A., Jovin, T.M., Hanau, L.H., Erlanger, B.F., and Wells, R.D. (1984) *J. Biol. Chem.* 259, 7268-7274.
34. Barker, W.C., and Dayhoff, M.O. (1977) *Comp. Biochem. Physiol.* 57B, 309-315.
35. Boguski, M.S., Elshourbagy, N., Taylor, J.M., and Gordon, J.I. (1985) *Proc. Natl. Acad. Sci. USA* 82, 992-996.
36. Karathanasis, S.K., Zannis, V.I., and Breslow, J.L. (1983) *Proc. Natl. Acad. Sci. USA* 80, 6147-6151.
37. Protter, A.A., Levy-Wilson, B., Miller, J., Bencen, G., White, T., and Seilhamer, J.J. (1984) *DNA* 3, 449-456.
38. Rees, A., Stocks, J., Shoulders, C.C., Galton, D.J., and Baralle, F.E. (1983) *Lancet* 444-446.