Volume 17 Number 2 1989 Nucleic Acids Research

Characterization of single base substitutions in edited apolipoprotein B transcripts

Gregory E.Tennyson, Charles A.Sabatos, Thomas L.Eggerman and H.Bryan Brewer Jr

Molecular Disease Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received September 23, 1988; Revised and Accepted December 7, 1988

ABSTRACT

Mature RNA transcripts from a single eukaryotic gene may contain different nucleotide sequences, ranging from alternately spliced exons to transcripts from separate alleles differing by only one base. Our laboratory and others have recently reported another class of RNA sequence differences, occurring in transcripts from the single copy apolipoprotein B (apoB) gene. A unique RNA editing mechanism allows expression of the CAA glutamine codon encoded by the apoB gene at nucleotide 6666, or terminates translation by the introduction of a premature UAA translational stop codon. In this study, we used the polymerase chain reaction (PCR) to amplify and characterize edited apoB RNA transcripts differing by a single nucleotide. Amplification and sequence analysis from small quantities of total RNA will facilitate the study of RNA editing and transcription in general.

INTRODUCTION

Two B apolipoproteins of molecular weights 512,000 and 250,000 designated apoB-100 and apoB-48, respectively, are present in human plasma (1). The production of these two proteins from the apoB gene is accomplished by a recently described RNA editing process (2-5). In the human liver and intestine (3), apoB-100 is encoded by a 14.1 kb mRNA, while apoB-48 is encoded by an edited mRNA containing a C to U substitution at nucleotide 6666, altering the glutamine (gln) codon (CAA) encoded by the gene, to a premature translational stop codon (UAA). The resultant mRNAs differ by a single base, yet the respective proteins, apoB-100 and apoB-48, differ significantly in size and function (1). Previous studies established the feasibility of amplification and qualitative analysis of RNA sequences, using the PCR with both the Klenow fragment and Taq polymerase (6-9). We now report amplification and quantitation of edited mRNAs differing by a single base.

MATERIAL AND METHODS

Enzymes and Chemicals:

Restriction enzymes were purchased from New England Biolabs. Avian myeloblastosis virus reverse transcriptase and human placental ribonuclease

inhibitor (RNAsin) were purchased from Promega. Thermophilus aquaticus (Taq) DNA polymerase was from Perkin-Elmer Cetus. T7 DNA polymerase (Sequenase) was purchased from United States Biochemicals. RNase A was from Sigma. The dNTPs were from Pharmacia, guanidine thiocyanate was from Fluka, and CsCl was purchased from Bethesda Research Laboratories.

NuSieve agarose was from FMC Bioproducts. α^{35}S dATP (1500 Ci/mmole) was from New England Nuclear. NA-45 paper was from Schleicher and Schuell. PCR primers were synthesized on an Applied Biosystems Model 380B DNA synthesizer.

Tissues and Cell Lines

Human liver samples were obtained from renal transplant donors or surgical biopsies as previously reported (10). Rat liver was from normal Sprague-Dawley adult males. Rabbit intestine was from normal New Zealand white males. The Caco-2 cell line, derived from human colon carcinoma cells, was obtained from American Type Culture Collection. All tissue samples were frozen in liquid N_2 immediately after excision. XL-1 Blue competent E. coli were from Stratagene.

Amplification of Edited RNAs

First Strand cDNA Synthesis: Ten micrograms of total RNA, isolated as previously described (11) from rat liver, rabbit intestine, human liver, and Caco-2 cells were incubated at 70°C for 10 minutes in the presence of 1 µg of antisense primer specific for each sequence (Fig. 1 and 2), 80 units of human placental ribonuclease inhibitor, 15 µl of 5X reverse transcriptase buffer (500 mM Tris/HCl pH 8.3 at 42°C, 50 mM $MgCl_2$, 700 mM KCl, 100 mM 2-mercaptoethanol), 150 µM dNTPs, and sterile H_2O in a final volume of 75 µl. The reaction mixture was allowed to cool to 42°C, 20 U of avian myeloblastosis virus reverse transcriptase added, and extension allowed to proceed at 42°C for 1 hour. The mixture containing first strand cDNA was phenol/chloroform extracted, ethanol precipitated, and resuspended in 50 µl of sterile H_2O.

PCR Amplification: Single stranded cDNAs (in 50 µl H_2O) were subjected to automated PCR amplification using the primers described. The reaction mixture (final volume 100 µl, with conditions as specified in the GeneAmp kit, Perkin Elmer Cetus) was incubated at 94°C for 6 minutes, followed by 30 cycles of successive denaturation (94°C for 30 seconds), annealing (55°C for 30 seconds), and extension (72°C for 2 minutes), followed by a final 10 minute extension at 72°C. 5µl of the PCR reaction mixture (5%) was withdrawn and 1 µl of RNase A (10 µg/µl) added; the mixture was incubated at 37°C for 10 minutes, and loaded directly on a 2% NuSieve - 1% agarose gel. Sequence Analysis of PCR Amplification Products: The remaining 95 µl of double stranded DNA containing-PCR product were incubated with RNase A (1 µl of RNase

A, 10 μg/μl) at 37°C for 30 minutes. The reaction mixture was phenol/chloroform extracted, ethanol precipitated, resuspended in TE, and digested overnight at 37°C with the appropriate restriction enzymes (for rat liver cDNA: BamHI and HindIII; for rabbit intestinal and Caco-2 cDNAs: EcoRI and BamHI). The digestion mixture was loaded directly onto a 1.8% agarose gel. Bands of predicted size (Fig. 1) were isolated on NA-45 paper, ligated to appropriate sites in M13mp19, and XL-1 Blue cells transformed. Clear, recombinant plaques were randomly selected (representing >90% of the total plaques) and sequenced using the dideoxy method (12). The sequencing reactions were resolved on 6% polyacrylamide, 7 M urea gels. Typically, 85-90% of the randomly selected plaques contained target apoB sequences for each RNA species analyzed.

RESULTS

Total RNAs were isolated from tissues of multiple species and the Caco-2 cell line. First strand cDNA was synthesized and amplified by the Taq PCR, using oligonucleotide primers flanking the region of interest (Figs. 1 and 2).

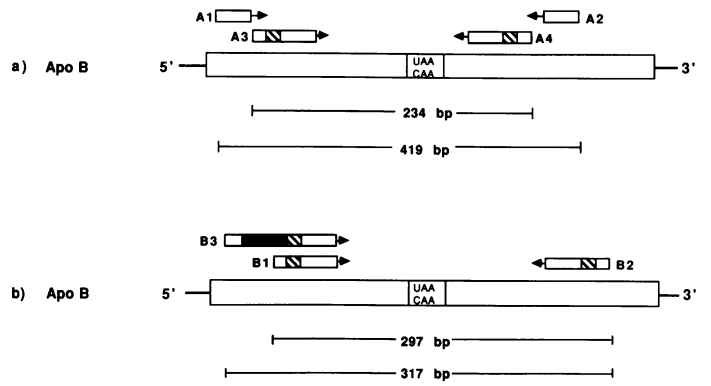

Fig. 1 Schematic representation of the apoB mRNA sequence flanking the RNA edited gln/stop codon. (a) The region containing the edited gln/stop codon in the apoB mRNAs is shown as an open box, with the location of the gln codon (CAA) and alternate translation termination codon (UAA) indicated. Primer pairs used to amplify each of the apoB mRNA species are: for human liver, A1 and A2, and A3 and A4; for rabbit intestine and Caco-2, A3 and A4. Relative locations of the primers flanking the edited codon are shown, with the direction of chain elongation shown by arrows. The open boxed region in each primer represents sequences which match the native sequence. The cross hatched region represents the portion of the primer containing the restriction enzyme site. Predicted sizes of amplified cDNAs are shown below. (b) For rat liver apoB mRNA amplification, primer pairs B1 and B2, and B3 and B2 were used. The solid boxed region in B3 represents the M13 sequencing primer.

A1	5'	CTGGGAAAACTCCCACAGCAA 3'
A2	5'	GTATCCACATTTTGAATCCA 3'
A3	5'	TTATCTGAaTtCATTCAATTGGGAGAGACAAGTT 3'
A4	5'	ATAATAggAtCcATAGCTATTTTcAAATCATGTA 3'
B1	5'	TAGAGgatCCCTGAGCAGGCTTCCTCAGCAG 3'
B2	5'	TTTAaagcTTCAATCATTCTATCAATAATCTG 3'
B3	5'	GATCAGTgTaaaAcGAcggccAGtGgatCCCTGAGCAGGCTTCCTCAGCAG 3'

Fig. 2 Specific apolipoprotein oligonucleotide primer sequences. The nucleotide sequences of the PCR primers illustrated in Fig. 1 are shown. Primers A1, A3, B1, and B3 are 5' sense primers; primers A2, A4, and B2 and 3' antisense primers. Restriction sites introduced by mismatches are underlined and include A3: EcoRI, A4: BamHI, B1: BamHI, B2: HindIII, B3: BamHI. The M13 sequence incorporated into the B3 primer is enclosed in the open box. Primer sequences which are identical to the native sequences are shown by upper case letters; mismatches are shown by lower case letters. Primers A3 and A4 are shown aligned with the rabbit apoB mRNA sequence. Primer A3 corresponds exactly with human apoB, while in primer A4, the c indicates a match (cytosine) with the human sequence.

Initially, human liver first strand cDNA was amplified with primers A1 and A2, to illustrate the feasibility of amplification of moderately abundant transcripts from small quantities of total RNA . A simple reverse transcription step followed by automated Taq PCR of the single stranded cDNA provided ample double stranded cDNA of the predicted size (419bp) (Fig. 3).

To facilitate directional cloning and sequence analysis of the amplified

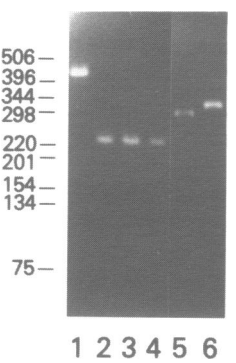

Fig. 3 Analysis of RNA amplification products. Agarose gel electrophoresis of apoB PCR products from amplification of human liver RNA (primers A1 and A2, lane 1; primers A3 and A4, lane 2), rabbit intestinal RNA (primers A3 and A4, lane 3), Caco-2 RNA (primers A3 and A4, lane 4), and rat liver RNA (primers B1 and B2, lane 5; primers B3 and B2, lane 6), separated through a 2% NuSieve, 1% agarose gel. Molecular sizes of a 1 kb ladder are shown on the left.

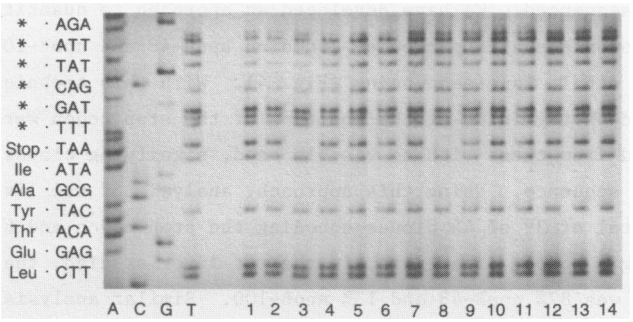

Fig. 4. Sequencing gel of rat liver apoB cDNA clones. A complete set of sequencing reactions from a stop codon containing clone from rat intestine is shown on the left. Predicted amino acids for the rat apoB-48 protein are included adjacent to the codons. The edited stop codon (TAA) is followed by asterisks. Lanes 1-14 contain single lane sequencing reactions from 14 rat liver cDNA clones using only dideoxythymidine terminators. Lanes 3 and 8 are from gln codon containing clones while the remaining lanes represent stop codon clones.

PCR products, we designed primers containing restriction enzyme sites (13), with native sequences mismatches (primer pairs A3 and A4, and B1 and B2, in Fig 2). Amplification of human liver, rabbit intestinal, and Caco-2 RNAs utilizing primer pair A3 and A4, as well as rat liver RNA with primer pair B1 and B2 yielded products of 234 bp, and 297 bp, respectively, as predicted (Figs. 1 and 3).

Finally, primers for RNA amplification incorporating both a restriction enzyme site as well as a sequencing primer (M13), containing significant native sequence mismatch, were designed (Figs 2 and 3). Amplification of rat liver RNA (primer pair B3 and B2) yielded a 317bp cDNA, the predicted size. Thus, the ability to amplify RNA as well as DNA (14) with PCR primers containing standard sequencing or phage promoter primers will allow direct PCR product sequencing with manual or automated sequencing (15). PCR products amplified from distinctly different RNAs may be sequenced with standard primers, eliminating the need for custom sequencing primers used in the triple primer method (16). The inclusion of a unique restriction site (not present in the amplified sequence) allows for removal of the standard sequencing primer site, should subcloning be undertaken.

For definitive analysis of the original apoB mRNA transcripts in the present study, the cDNAs obtained from PCR amplification of rat liver, rabbit intestine, and Caco-2 RNAs were directionally cloned into M13mp19, using the incorporated restriction sites (Fig. 2). Recombinant plaques were randomly

selected and sequenced. We have developed an approach to quantitate the ratio of stop:gln codon containing clones, encoding apoB-48 and apoB-100, respectively, using single lane sequencing (Fig. 4). With this technique, clones containing a deoxythymidine at the position of the stop codon were easily differentiated from those with no visible band, signifying clones containing the gln codon sequence. Using this approach, analysis of rat liver RNA revealed a total of 39 of 45 clones encoding the stop codon and 6 of 45 encoding the gln codon. Therefore, in the rat liver studied, the percentage of apoB mRNAs was 87% apoB-48 and 13% apoB-100. Similar analysis showed the rabbit intestine to contain 83% apoB-48 and 17% apoB-100, while Caco-2 cells contained 100% apoB-100 in the original mRNAs, under the conditions studied.

DISCUSSION

Two colinear B apolipoproteins, apoB-48 and apoB-100, circulate in plasma and differ remarkably in function (1): apoB-100 serves as the ligand for binding of low density lipoproteins (LDL) to the LDL receptor, whereas the specific function(s) of apoB-48, the amino-terminal 1-2152 residues of apoB-100, remains unknown. Recent studies have established a novel RNA editing mechanism which directs production of two different protein isoforms of apoB from the single copy gene (2-5). In this paper, we have used the PCR to evaluate the presence of RNA editing of apoB in the rat liver, rabbit intestine, and in Caco-2 cell, an intestinal cell line used as a model for intestinal apolipoprotein biosynthesis (17). Based on the nucleotide sequences for human and rabbit apoB, and the sequence for the edited codon flanking region of apoB recently determined in this laboratory for the rat, we designed primers flanking the gln/stop codon for use in the PCR. Using this technique, we found the ratio of apoB-48 mRNA to apoB-100 mRNA to be greater in the rat liver and rabbit intestine than in Caco-2 cells. In summary, the combined data from these studies indicate a marked contrast in the degree of RNA editing in different tissues and species, resulting in significant differences in the apoB isoproteins synthesized by the individual tissues.

Although we can not rigorously exclude the possibility that the PCR primers designed in this study might selectively amplify either the edited or encoded transcripts, differing by one base, we believe this to be unlikely. Firstly, the PCR primers hybridize some distance from the site subject to base substitution. Secondly, computerized secondary structure analysis [RNAFOLD, (18)] of apoB RNAs in the region immediately flanking the stop/gln codon shows stop and gln codon species to model identically (not shown). Finally, results

of analysis of Caco-2 RNAs using the techniques outlined herein clearly showed predominance of the encoded (CAA) apoB species (100%) in Caco-2 cells, a finding consistent with the protein data from Lee et al. (19), showing only apoB-100 secretion in this cell line by SDS-PAGE. Similarly, the rat liver was predicted to contain predominantly the edited (TAA) apoB transcript (87%), encoding the apoB-48 isoprotein,consistent with the published protein data of Sparks et al. (20) and Coleman et al. (21) showing a predominance of the apoB-48 isoprotein in the rat liver. While PCR amplification of RNA species such as apoB as described herein may not be absolutely quantitative, these observations suggest the method to be quite useful in identifying the predominant species.

The PCR has proven to be an important tool for amplification and analysis of edited RNA transcripts. In addition, traditional cDNA library construction, screening, and subcloning requires considerably more time, effort, and template RNA than does PCR amplification and analysis of RNA. Direct sequencing of small regions of RNA is feasible (22), but quantitation of transcripts has not been described with direct sequencing. Further applications of RNA amplification in RNA sequence analysis are forseeable. RNA transcripts encoded by genes differing by a nucleotide or nucleotides creating or abolishing restriction sites may be amenable to amplification, digestion and analysis as reported for genomic DNA (23). The use of PCR primers which flank small, alternately spliced exons, such as thyroid hormone receptor transcripts (24,25), may allow the determination of the ratios of spliced products. Quantitation of the prevalence of transcripts from very similar alleles or genes may also be possible.

In this report, we have extended the utility of PCR amplification to include detection of single base substitutions created by RNA editing, using small amounts of total RNA as template. Reverse transcription provides initial amplification through first strand cDNA production. Taq PCR amplification of these cDNAs yields ample double stranded cDNAs for directional subcloning, sequencing, and estimation of the amount of editing in the original mRNAs. Analysis of more extensively edited RNAs, such as those recently described in Trypanosoma brucei and other protozoans (26,27) using the method described here may facilitate study of more widespread editing, although editing of apoB and T. brucei transcripts may occur through two completely different mechanisms. Amplification and analysis of gene transcripts using the PCR should add valuable information regarding RNA editing as well as the process of transcription in general.

ACKNOWLEDGEMENTS

We thank Joanie Gault for the careful preparation of this manuscript.

REFERENCES

1. Kane, J.P. (1983) Ann. Rev. Physiol. 45, 637-650.
2. Hospattankar, A.V., Higuchi, K., Law, S.W., Meglin, N.M. and Brewer, H.B., Jr. (1987) Biochem. Biophys. Res. Commun. 148, 279-285.
3. Powell, L.M., Wallis, S.C., Pease, R.J., Edwards, Y.H., Knott, T.J. and Scott, J. (1987) Cell 50, 831-840.
4. Chen, S.H., Habib, G., Yang, C.Y., et al. (1987) Science 238, 363-366.
5. Higuchi, K., Hospattankar, A.V., Law, S.W., Meglin, N., Cortright, J. and Brewer, H.B., Jr. (1988) Proc. Natl. Acad. Sci., U.S.A. 85, 1772-1776.
6. Todd, J.A., Bell, J.I. and McDevitt, H.O. (1987) Nature 329, 599-604.
7. Lee, C.L., Wu, X., Gibbs, R.A., Cook, R.G., Muzny, D.M. and Caskey, C.T. (1988) Science 239, 1288-1291.
8. Kawasaki, E.S., Clark, S.S., Coyne, M.Y., Smith, S.D., Champlin, R., Witte, O.N. and McCormick, F.P. (1988) Proc. Natl. Acad. Sci. USA 85, 5698-5702.
9. Sarkar, G. and Sommer, S.S. (1988) Nucl. Acid Res. 16, 5197.
10. Edge, S.B., Hoeg, J.M., Triche, T., Schneider, P.D., and Brewer, H.B., Jr. (1986). J. Biol. Chem. 261, 3800-3806.
11. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
12. Sanger, F., Nickeln, S. and Coulson, A.R. (1977) Proc. Natl. Acad.Sci., U.S.A. 74, 5463-5467.
13. Scharf, S.J., Horn, G.T. and Erlich, H.A. (1986) Science 233, 1076-1078.
14. Stoflet, E.S., Koerbel, D.D., Sarkar, G. and Sommer, S.S. (1988) Science 239, 491-494.
15. Wong, C., Dowling, C.E., Saiki, R.K., Higuchi, R.G., Erlich, H.A. and Kazazian, H.H. (1987) Nature 330, 384-386.
16. Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. and Wilson, A.C. (1987) Nucleic Acids Res. 15, 529-542.
17. Hughes, T.E., Sasak, W.V., Ordovas, J.M., Forte, T.M., Lamon-Fava, S., and Schaefer, E.J. (1987) J. Biol. Chem. 262, 3762-3767.
18. RNAFOLD (Ver. 1.06, 1987); PCGENE: Intelligentics, Inc., Mountainview, CA.
19. Lee, D.M., Dashti, N., Mok, T. (1988) Biochem. Biophys. Res. Commun. 156, 581-587.
20. Sparks, J.D., Sparks, C.E., Bolognino, M., Roncone, A.M., Jackson, T.K., Amatruda, J.M. (1988) J. CLin. Invest. 82, 37-43.
21. Coleman, R.A., Haynes, E.B., Sand, T.M., Davis, R.A. (1988) J. Lipid Res. 29, 33-51.
22. Feagin, J.E., Jasmer, D.J. and Stuart, K. (1987) Cell 49, 337-345.
23. Kogan, S.C., Doherty, M. and Gitscher, J. (1988) N. Engl. J. Med. 317, 985-990.
24. Mitsuhashi, T., Tennyson, G. and Nikodem, V. (1988) Nucl. Acid Res. 16, 5697.
25. Mitsuhashi,T., Tennyson, G.E. and Nikodem, V.M. (1988) Proc. Natl. Acad. Sci., U.S.A. 85, 5804-5808.
26. Shaw, J.M., Feagin, J.E., Stuart, K. and Simpson, L. (1988) Cell 53, 401-411.
27. Feagin, J.E., Abraham, J.M. and Stuart, K. (1988) Cell 53, 413-422.