

Detection of single base differences using biotinylated nucleotides with very long linker arms

Kenneth J. Livak, Frank W. Hobbs and Robert J. Zagursky

The Du Pont Merck Pharmaceutical Company, Experimental Station, PO Box 80328, Wilmington, DE 19880-0328, USA

Received June 8, 1992; Revised and Accepted August 12, 1992

ABSTRACT

A simple primer extension method for detecting nucleotide differences is based on the substitution of mobility-shifting analogs for natural nucleotides (1). This technique can detect any single-base difference that might occur including previously unknown mutations or polymorphisms. Two technical limitations of the original procedure have now been addressed. First, switching to *Thermococcus litoralis* DNA polymerase has eliminated variability believed to be due to the addition of an extra, non-templated base to the 3' end of DNA by *Taq* DNA polymerase. Second, with the analogs used in the original study, the mobility shift induced by a single base change can usually be resolved only in DNA segments 200 nt or smaller. This size limitation has been overcome by synthesizing biotinylated nucleotides with extraordinarily long linker arms (36 atom backbone). Using these new analogs and conventional sequencing gels (0.4 mm thick), mutations in the human β -hexosaminidase α and *CYP2D6* genes have been detected in DNA segments up to 300 nt in length. By using very thin (0.15 mm) gels, single-base polymorphisms in the human *APOE* gene have been detected in 500-nt segments.

INTRODUCTION

The basis of molecular genetics is the identification and characterization of mutations. Thus, many techniques have been developed to compare homologous segments of DNA to determine if the segments are identical or different. We have described a simple primer extension assay that distinguishes homologous DNA segments differing by as little as a single nucleotide (1). DNA strands are synthesized with one of the four natural nucleotides replaced with an analog that retards electrophoretic mobility ('mobility-shifting' analog). For example, incorporation of biotin-11-dUTP, a commercially available analog of TTP, into a DNA strand causes a one nucleotide mobility shift when the DNA is fractionated on a sequencing gel. This means that, for each biotinylated residue incorporated, the biotin-containing DNA strand migrates at a position approximately one nucleotide slower than is expected based on the length of the DNA strand. DNAs that are the same length but differ in the number of analog molecules per strand migrate differently on a sequencing gel and thus are distinguished.

The assay is sensitive—by using two analogs (in separate reactions) and analyzing both strands, any single nucleotide change can be detected. Furthermore, the assay can identify previously undetected mutations. Another advantage is that the method uses conventional molecular biology techniques. The primer extension reactions and gels are basically the same as those used in standard dideoxynucleotide sequencing.

The ability to use mobility-shifting nucleotide analogs to detect single base differences was first demonstrated by examining a segment of the human insulin receptor gene (1). When the method was used to examine a number of additional DNA segments, two limitations of the technique became apparent. First, using the nucleotide analogs described in Kornher and Livak (1), the method best distinguishes single base differences when the DNA segments examined are 200 nt or less. Increasing the size limit would greatly increase the utility of this technique. Second, there was run-to-run variation in the appearance of shadow bands migrating just faster than the main bands. At its worst, the shadow bands made it difficult to distinguish homozygotes, which should exhibit one prominent band, from heterozygotes, for which two prominent bands are expected. As described in this report, these two problems were overcome by using a different enzyme to perform the primer extension reactions, by synthesizing new nucleotide analogs with extraordinarily long linker arms, and by using very thin polyacrylamide gels to reduce electrophoresis time.

MATERIALS AND METHODS

Synthesis of nucleotide analogs

Biotin-36-dUTP (1b). Biotin-XX-NHS ester (22.7 mg, 40 μ mol, Clontech #5002) in 200 μ l of dimethylformamide was heated with occasional vortexing at 50°C until all of the solid dissolved. A solution of the triethylammonium salt of dUTP-36 (1a, 20 μ mol, 242 ODU at 290 nm; 2) and N-methylmorpholine (22 μ l, 200 μ mol) in water (133 μ l) and dimethylformamide (67 μ l) was added. The reaction was allowed to proceed with occasional vortexing at 25°C for 4 h. The reaction was filtered, diluted with 5 ml of 0.1 M TEAB (aqueous triethylammonium bicarbonate), and loaded onto a DEAE-A25—120 Sephadex column (2.6 \times 40 cm, prewashed with 1.0 M TEAB and equilibrated with 3 column volumes of 0.1 M TEAB). The column was eluted with a linear gradient from 0.1 M (200 ml) to 1.0 M (200 ml) TEAB at a flow of 2 ml/min while monitoring UV absorption at 260 nm.

(CO₂ was continuously bubbled into the TEAB.) The major peak eluting from 250 to 290 ml was collected, concentrated and co-evaporated twice with ethanol to afford 242 ODU (77%) of white foam. The material was further purified by preparative HPLC (1×25 cm ODS column, 5–35% acetonitrile in 0.1 M TEAB (pH = 7) gradient over 30 min with a flow rate of 3 ml/min). The major peak was collected, co-evaporated and lyophilized. The product was dissolved in water (1 ml) containing triethylamine (5 µl) and eluted through a column (1 ml) of the sodium form of Dowex AG 50-X8 with water. Fractions (ca. 1 ml each) containing UV absorbing material were combined and lyophilized to afford biotin-36-dUTP, **1b**, (153 ODU at 290 nm, 12.6 µmol, 63%) as a white foam. UV (water): 212, 235 and 290 nm. (The extinction coefficients were 16,000, 12,400 and 12,100 based on an extinction coefficient of 12,100 at 290 nm previously established for this chromophore.) ³¹P-NMR: δ = -9.2, -9.7, and -21.4. ¹H-NMR: δ = 8.15 (s, 1H, H6), 6.27 (t, 1H, H1'), 4.65 (m, 2H, H3' and H7'), 4.52 (s, 2H, CCCH₂O), 4.45 (dd, 1H, H8"), 4.20 (m, 3H, H4' and H5'), 3.6–3.9 (m, 22 H, OCH₂CH₂O), 3.42 (t, 2H, OCH₂CH₂N), 3.35 (dd, 1H, H6"), 3.20 (m, 4H, CH₂N), 3.00 (dd, 1H, H9'a), 2.77 (d, 1H, H9'b), 2.2–2.5 (m, 8H, H2' and CH₂C=O) and 1.3–1.8 (m, 18H, CH₂). (Carbons on the biotin subunit of the molecule were assigned double-prime numbers beginning with the carboxylic acid carbon.)

Biotin-36-dCTP (1d). The triethylammonium salt of dCTP-36 (**1c**, 20 µmol, 186 ODU at 294 nm; 2) was biotinylated and purified as described above to afford biotin-36-dCTP, **1d**, (121 ODU at 294 nm, 13.4 µmol, 65%) as a white foam. UV (water): 213, 238, and 294 nm. (The extinction coefficients are 28,000, 15,000 and 9,300 based on an extinction coefficient of 9,300 at 294 nm previously established for this chromophore.) ³¹P-NMR: δ = -5.60 (br s, 1P), -9.51 (d, 1P) and -20.63 (t, 1P). ¹H-NMR: δ = 8.22 (s, 1H, H6), 6.28 (t, 1H, H1'), 4.65 (m, 2H, H3' and H7'), 4.52 (s, 2H, CCCH₂O), 4.45 (dd, 1H, H8"), 4.24 (m, 3H, H4' and H5'), 3.6–3.9 (m, 22 H, OCH₂CH₂O), 3.42 (t, 2H, OCH₂CH₂N), 3.35 (dd, 1H, H6"), 3.20 (m, 4H, CH₂N), 3.03 (dd, 1H, H9'a), 2.81 (d, 1H, H9'b), 2.2–2.5 (m, 8H, H2' and CH₂C=O), 1.94 (s, acetate impurity), and 1.3–1.8 (m, 18H, CH₂).

Biotin-36-dc⁷ATP (2b). The triethylammonium salt of dc⁷ATP-36 (**2a**, 10 µmol, 127 ODU at 280 nm; 2) was biotinylated and purified as described above to afford biotin-36-dc⁷ATP, **2b**, (25.4 ODU at 280 nm, 2.0 µmol, 20%) as a white foam. The following changes were made: a) The reaction was run on one-half the above scale. b) A 1×19 cm ion exchange column was used and elution was with 150 ml of each buffer. c) From preparative HPLC, two equal fractions were collected and the faster eluting fraction was repurified by preparative HPLC. UV (water): 214, 239 and 280 nm. (The extinction coefficients were 20,300, 13,100 and 12,700 based on an extinction coefficient of 12,700 at 280 nm previously established for this chromophore.)

PCR amplifications

The buffers used in PCR amplifications were *Taq* buffer (10 mM Tris-HCl, pH 8.3; 50 mM KCl; 5 µM EDTA; 0.01% gelatin) or Stoffel buffer (10 mM Tris-HCl, pH 8.3; 10 mM KCl). All reactions contained 100 ng human genomic DNA and 200 µM

each dNTP, and were performed in 50 µl reactions under mineral oil in a DNA Thermal Cycler (Perkin-Elmer Cetus).

The primers used to amplify 113 bp and 498 bp segments of the β-hexosaminidase α gene were: TS1-GTGTGGCGAGAG-GATATTCAGT (exon 11 primer from Myerowitz and Costigan (3)); TS6-TCCTGCTCTCAGGCCCAACCCTC (intron 12 primer PCA₂ from Myerowitz (4)); TS10-AAGCT-TCACTCTGAGCATAACAAG (intron 7 primer from Navon and Proia (5)); and TS14-TATCCGTGTGCTTGCGAGGTTTG (nucleotides 749–771 in Myerowitz *et al.* (6)). Reactions contained *Taq* buffer, 2.5 mM MgCl₂, 200 nM each TS primer, plus 1 unit *AmpliTaq* polymerase (Perkin-Elmer Cetus) and were subjected to the following regimen: 92°C (3 min); 30 cycles of ramp to 92°C (20 sec), 92°C (40 sec), ramp to 68°C (30 sec), 68°C (1 min, increase by 3 sec each cycle); 72°C (5 min).

The primers used to amplify a 297 bp segment of the *CYP2D6* gene were: CY11-CCGCCTTCGCCAACCCT (nucleotides 1826–1843 in Kimura *et al.* (7)); and CY2-GAGACTCCTCG-GTCTCTC (primer 2 from Heim and Meyer (8)). Reactions contained Stoffel buffer, 5 mM MgCl₂, 1 µM each CY primer, plus 5 units Stoffel fragment (Perkin-Elmer Cetus) and were subjected to the following regimen: 98°C (3 min); 5 cycles of 98°C (1 min), 62°C (1 min), 72°C (1 min); 30 cycles of 95°C (1 min), 62°C (1 min), 72°C (1 min); 72°C (5 min).

The primers used to amplify a 538 bp segment of the *APOE* gene were: ApoE11-GGCACGGCTGTCCAAGG (adapted from primer F6 of Emi *et al.* (9)); and ApoE24-CACCAGGGGCT-CGAACC (nucleotides 4218–4202 in Paik *et al.* (10)). Reactions contained *Taq* buffer, 1.5 mM MgCl₂, 10% DMSO, 1 µM each ApoE primer, plus 4 units *AmpliTaq* polymerase and were subjected to the following regimen: 95°C (3 min); 35 cycles of 95°C (1 min), 58°C (1 min), 70°C (1 min); 70°C (5 min).

After amplification, each DNA sample was transferred to a fresh tube, treated with 1 unit calf intestinal alkaline phosphatase (Promega) at 37°C for 15 min to hydrolyze residual dNTPs, and precipitated by adding 50 µl 5 M NH₄-acetate and 100 µl isopropanol. After sitting at room temperature for 10 min, the DNA is collected by a 10 min centrifugation, rinsed with 200 µl cold 70% ethanol, dried, dissolved in 20 µl TE-8.0 (10 mM Tris-HCl, pH 8.0; 1 mM EDTA), incubated at 70°C for 15 min, and stored at 4°C. This procedure eliminates the dNTPs and most of the primers from the PCR reaction.

Primer extensions with nucleotide analogs

For *n* samples, a Master Mix was prepared containing (n+1)×1 µl H₂O; (n+1)×0.5 µl 10×*Tli* buffer [200 mM Tris-HCl, pH 8.8; 100 mM (NH₄)₂SO₄; 100 mM KCl; 20 mM MgSO₄; 1% Triton X-100]; (n+1)×0.5 µl DMSO; (n+1)×0.1 µl *Thermococcus litoralis* (*Tli*) DNA polymerase (Vent DNA polymerase, 1 U/µl; New England BioLabs); and (n+1)×0.9 µl labeled primer. For each primer extension reaction, 1 µl analog mix (1 mM biotin-36-dATP, -dCTP, or -dUTP plus 500 µM of each of the remaining three dNTPs) and 1 µl PCR-amplified template DNA were put in an Eppendorf tube. After adding 3 µl Master Mix to each reaction, the tubes were incubated in a boiling water bath for 1 min, at room temperature for 1 min, and at 75°C for 3 min. For ApoE24 primer extensions this cycle was repeated four additional times. Reactions were terminated by adding 10 µl formamide-dye solution (95% formamide; 12.5 mM EDTA; 0.3% bromophenol blue; 0.3% xylene cyanol) and incubating in a boiling water bath for 3 min.

The labeled primers used to analyze the Tay Sachs mutations in the β -hexosaminidase a gene were: TS15-TCTGGTCCCAG-ACATCATTC (intron 7 sequence from footnote 31 in Navon and Proia (5)); TS16-CCTTCCAGTCAGGGCCATA (nucleotides 1297–1279 in Myerowitz *et al.* (6)); and TS17-TTGGTGGAGAGGCTTGATG (nucleotides 1359–1378 in Myerowitz *et al.* (6)). These TS primers were synthesized with a succinyl fluoroscein dye (SF-505; 11) attached to their 5' ends (12). The final concentrations of the primers in the extension reaction were 100 nM TS15, 300 nM TS16, and 100 nM TS17. Reactions with the fluorescent primers were terminated by the addition of 10 μ l GENESIS loading solution (Du Pont) rather than the formamide-dye solution. The labeled primers used to analyze the *CYP2D6* and *APOE* genes were: CY2 (PCR primer); ApoE24 (PCR primer); and ApoE13-GCGCGGACATGGAG-GAC (nucleotides 3725–3741 in Paik *et al.* (10)). These primers were labeled with 32 P by mixing 7 μ l (70 μ Ci) γ - 32 P-ATP, 1 μ l 10 \times kinase buffer (500 mM Tris-HCl, pH 7.5; 100 mM MgCl₂; 50 mM dithiothreitol; 1 mM spermidine), 1 μ l 10 μ M primer, 1 μ l (10 units) polynucleotide kinase (New England BioLabs) and incubating at 37°C for 10 min. The final concentration of each primer in an extension reaction was 18 nM CY2, 180 nM ApoE24, or 36 nM ApoE13.

Gel electrophoresis

The buffer used to run the sequencing gels was a modified form of TBE designated MTB ($10 \times \text{MTB} = 163.5 \text{ g Tris base, } 27.8 \text{ g boric acid, } 9.3 \text{ g Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O per 1 liter}$). Except for the C analog gel in Figure 4, electrophoresis was performed with $1 \times \text{MTB}$ in both the gel and running buffer. Gels were 6% 19:1 acrylamide/bis-acrylamide and contained either 7 M or 8 M urea (7 M urea seemed to give slightly better separation). Very thin gels were approximately 0.15 mm thick, 32 cm long, and 20 cm wide. They were poured from the bottom by capillary action using adhesive tape (Serva cat. #42927) as side spacers and 0.2 mm combs (IBI) wedged in the top. At first, the very thin gels were transferred to 3MM paper (Whatman) after electrophoresis and dried down. At times, though, the bands were not perfectly straight. The waviness observed in bands can be seen in the 0.15 mm gel of Figure 3. Thus, for the analysis of 500 nt primer extension products in Figure 4, the gels were covalently attached to the glass by treating the larger plate with γ -methacryloxypropyl-trimethoxy silane (Sigma) as described by Garoff and Ansorge (13). After electrophoresis, the gel was covered with Reynolds 904 plastic film and dried directly onto the glass plate under vacuum.

RESULTS

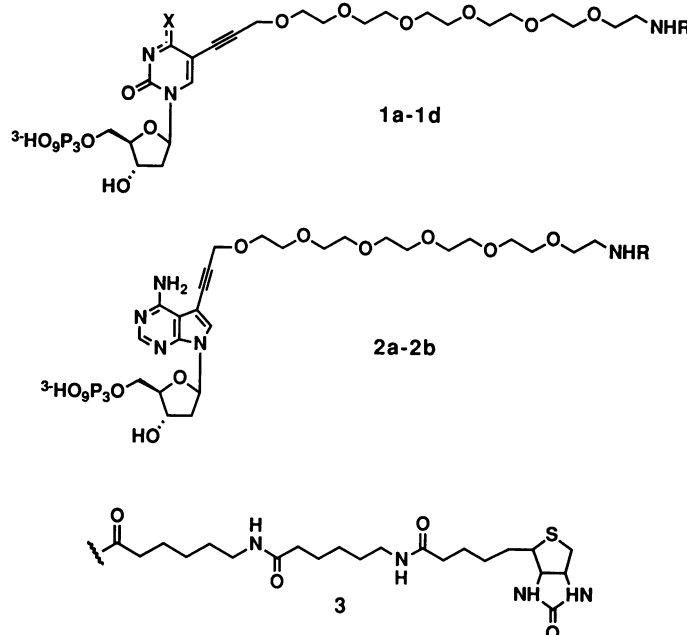
DNA polymerase used in the assay

The mobility-shift method for polymorphism detection was originally developed using *Taq* DNA polymerase to incorporate nucleotide analogs as part of a primer extension reaction. When these primer extension products were examined on a denaturing polyacrylamide gels, shadow bands were sometimes seen migrating just faster than the main bands. The appearance of these extra bands was finally traced to the ability of *Taq* DNA polymerase to add a single, non-templated nucleotide, usually A, to the 3' end of blunt-end DNA fragments (14). In fact, the main bands observed on gels contained this extra base; whereas,

shadow bands were the result of failing to add the non-templated base. Variation was observed because the extent of extra nucleotide addition depends on enzyme lot, enzyme concentration, nucleotide concentration, exact reaction conditions, and possibly sequence context. This problem was overcome by switching to a different DNA polymerase. In order to retain the convenience and sensitivity of using a thermostable enzyme, the polymerase chosen was *Tli* DNA polymerase. *Tli* DNA polymerase possesses 3' exonuclease activity and thus does not leave an extra, non-templated nucleotide on DNA fragments. The use of *Tli* DNA polymerase greatly reduces problems in interpretation due to shadow bands.

Nucleotide analogs with long linker arms

The switch to *Tli* DNA polymerase led to another fortunate result. A series of nucleotide analogs with longer linker arms had been synthesized and tested for acceptance by *Taq* DNA polymerase. Figure 1 shows three analogs synthesized with a linker of 36 atoms, the longest linker so far constructed. The analogs of Figure 1 were not accepted that well by *Taq* DNA polymerase, but were found to be fairly good substrates for *Tli* DNA polymerase. Using these analogs with *Tli* DNA polymerase dramatically improved the results obtained with the mobility-shift technique for polymorphism detection.



- 1a** X = O, R = H₂⁺. dUTP-22.
1b X = O, R = 3. Biotin-36-dUTP.
1c X = NH₂, R = H₂⁺. dCTP-22.
1d X = NH₂, R = 3. Biotin-36-dCTP.
2a X = O, R = H₂⁺. dc⁷ATP-22.
2b X = O, R = 3. Biotin-36-dc⁷ATP.

Figure 1. Biotinylated nucleotides with very long linker arms. Structures are shown for the intermediates and final compounds used in the mobility shift assay.

Tay Sachs mutations

Long linker analogs were first used in a nonradioactive assay for detecting mutations implicated in Tay Sachs disease. Figure 2A shows the segments of the β -hexosaminidase α gene analyzed and the fluorescent primers that were used to assay for each specific mutation. The 113-bp and 498-bp segments were amplified from human genomic DNA in a single reaction. After elimination of the dNTPs used in PCR, the amplified segments were used as templates in a primer extension reaction containing the three fluorescent primers, biotin-36-dCTP, dATP, dGTP, TTP, and *Tli* DNA polymerase. Figure 2B shows fluorescent primer extension products electrophoresed through a denaturing polyacrylamide gel and detected using the GENESIS 2000 system. The normal individual shows a single peak for each exon because this person is homozygous normal at all three mutation sites. Heterozygotes are detected by the appearance of two peaks. For exon 7, the peak for the mutant allele runs faster than the normal peak because the mutant primer extension product has one less biotinylated C residue. The mutant peak for exon 12 runs slower because this product has one additional C analog; and the mutant peak for exon 11 runs slower because this product has four additional non-analog nucleotides. Figure 2B clearly shows that heterozygotes can be detected for each of the three exons. Furthermore, this analysis demonstrates that more than one mutation can be assayed in a single sample by using multiplex PCR and multiple primers in the extension reaction. The only requirement is that the primers used produce extension products migrating at different points in the gel. The results of Figure 2B also show that the use of fluorescent primers provides a convenient method for achieving nonradioactive detection in the mobility shift assay.

CYP2D6 mutation

The use of the biotin-36 analogs also makes it possible to analyze primer extension products longer than 200 nt. Figure 3 presents an analysis using extension products of approximately 300 nt performed on a segment of the human *CYP2D6* gene. *CYP2D6* encodes the cytochrome P450 debrisoquine 4-hydroxylase. Mutations in this gene cause poor metabolism of debrisoquine and over 25 other drugs, a phenotype inherited as an autosomal recessive trait by 5–10% of Caucasians. One of the most common mutations that inactivates the *CYP2D6* gene is a G-to-A change that eliminates the splice acceptor site at the intron 3/exon 4 boundary (15). As diagrammed in Figure 3, primer extensions across this mutation site were performed using either a C or a T analog. Detection of the *CYP2D6* mutation is shown by the three individuals who exhibit two bands. These people are heterozygous for the mutation and thus are carriers for the 'poor metabolizer' phenotype. The other three individuals have only one of the two bands, and this is the slower band with the C analog and the faster band with the T analog. This electrophoretic behavior indicates there is a C at the mutation site and therefore the individuals with a single band are homozygous for the normal allele. The key point is that single base differences were detected in primer extension products of 297 nt. This is nearly 100 nt longer than the products that could be analyzed using analogs with shorter linker arms. In analyzing these longer extension products, the time required for electrophoresis becomes a significant factor. Figure 3 shows detection of single base differences when the products were analyzed on a conventional 0.4 mm sequencing gel. This required a 7 h electrophoresis run. By using a very thin (0.15 mm) gel,

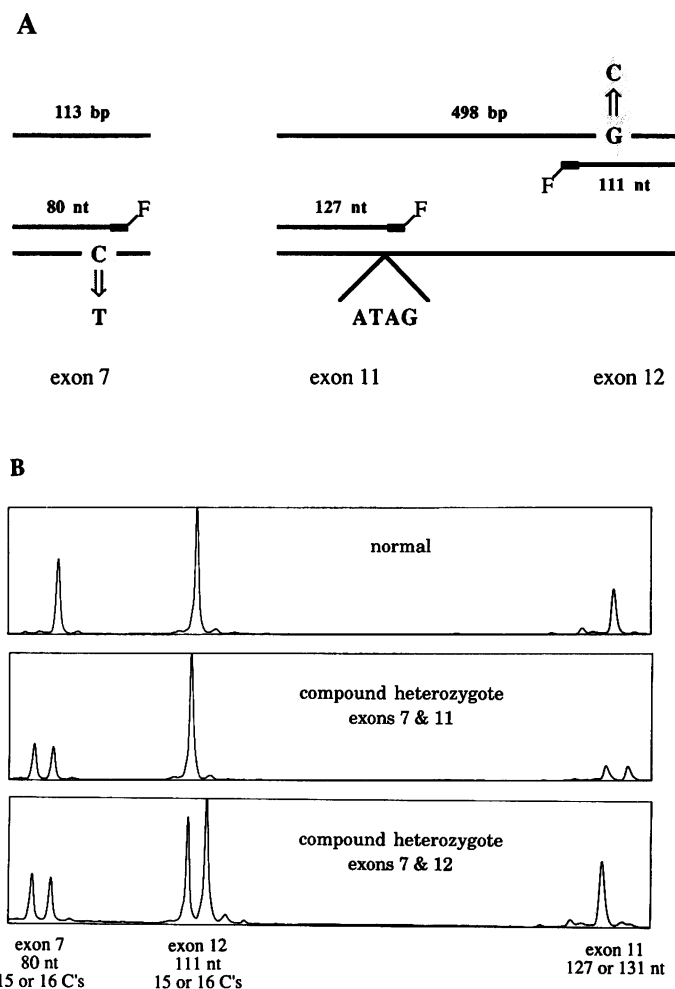


Figure 2. Detection of Tay Sachs mutations using fluorescence-labeled primers. A: Scheme for amplifying two segments (113 bp and 498 bp) of the β -hexosaminidase α gene and detecting three Tay Sachs mutations (3–5) using the mobility shift assay. B: As described in Materials and Methods, DNA segments containing the Tay Sachs mutation sites were amplified from three individuals—one normal and two compound heterozygotes. These PCR-amplified segments were used as templates for the primer extension reaction diagrammed in A. Primer extension products (80 nt, 127 nt, and 111 nt) incorporating biotin-36-dCTP were synthesized in a single reaction using the fluorescence-tagged (F) primers TS15 (exon 7), TS16 (exon 11), and TS17 (exon 12). The extension products were electrophoresed on a 6% acrylamide/8 M urea gel at 24 watts in the GENESIS 2000 DNA Analysis System (Du Pont) following the manufacturer's protocol. Output from the GENESIS 2000 is shown as fluorescence intensity plotted versus electrophoresis time. As indicated, the exon 7 and exon 11 products each contained 15 or 16 biotinylated C residues.

approximately the same separation of bands was achieved in 4 h. Thus, the use of 0.15 mm gels enhances the convenience of the mobility-shift technique. One final point about the *CYP2D6* results is that the CY2 primer was used both to amplify the *CYP2D6* segment and as the labeled primer in the extension reactions. This demonstrates that a nested primer is not required for mobility-shift analysis as long as the PCR products do not have excessive background.

Apo E polymorphisms

The *APOE* gene has three common alleles $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$, which were originally detected by isoelectric focusing of the apoE protein (reviewed in 16). Sequencing has demonstrated that the

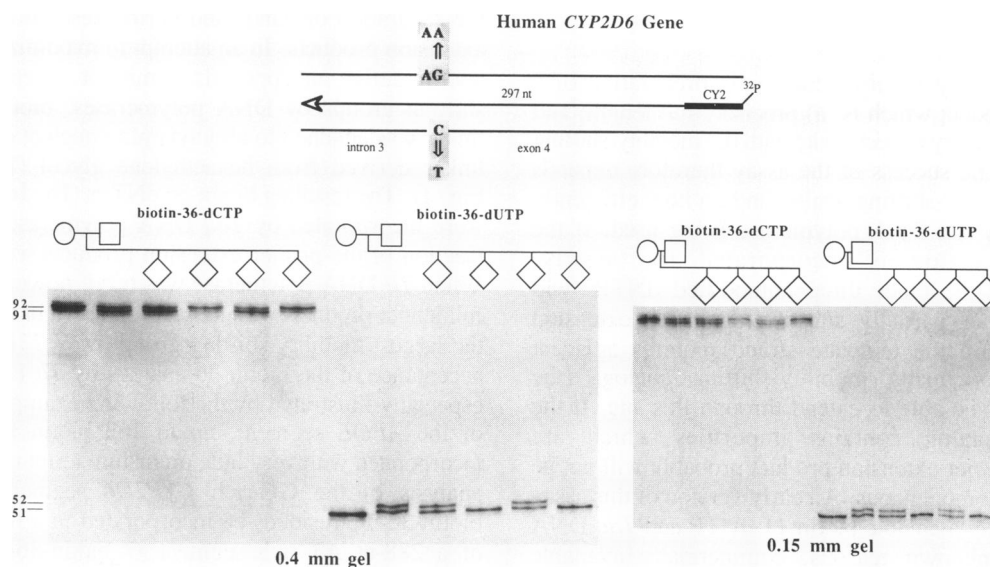


Figure 3. Detection of a mutation in the human *CYP2D6* gene using primer extension products of about 300 nt. Primers CY11 and CY2 were used to amplify a 297-nt segment of the *CYP2D6* gene from the parents and 4 children of CEPH pedigree 1333. These segments were used as templates in primer extension reactions containing ^{32}P -labeled CY2 and either biotin-36-dCTP or biotin-36-dUTP as the mobility-shifting analog. The left panel shows the products electrophoresed on a 6% acrylamide/8 M urea 0.4 mm x 40 cm gel run at 90 watts for 7 h. The right panel shows the products electrophoresed on a 6% acrylamide/7 M urea 0.15 mm x 32 cm gel run at 40 watts for 4 h. After electrophoresis, the gels were transferred to 3MM paper (Whatman), dried under vacuum, and exposed to X-ray film. The numbers to the left show the number of modified dC (91 or 92) or dU (51 or 52) residues in the bands. All extension products are 297 nt in length, but the C analog products run slower because they contain approximately 40 additional biotinylated residues.

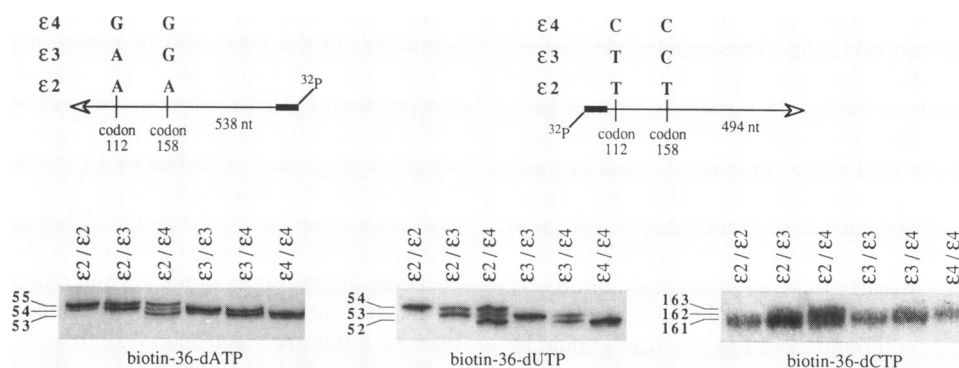


Figure 4. Detection of polymorphisms in the human *APOE* gene using primer extension products of about 500 nt. Out of a screen of *APOE* genotypes, an individual was selected to represent each of the six possible genotypes. Primers ApoE11 and ApoE24 were used to amplify a 538-nt segment of the *APOE* gene from these six individuals. Primer extensions incorporating biotin-36-dATP, -dUTP, or -dCTP were performed across the polymorphic sites as diagrammed. The A and T analog products were electrophoresed on 6% acrylamide/7 M urea/1 x MTB 0.15 mm x 32 cm gels using 1 x MTB as running buffer. The gels were run at 40 watts for 6 h, then dried onto the larger glass plate and exposed to X-ray film. The C analog products were electrophoresed on a 5% acrylamide/7 M urea/1.2 x MTB 0.15 mm x 32 cm gel using 0.6 x MTB as running buffer. The gel was run at 40 watts for 7 h, then dried onto the larger glass plate and exposed to a storage phosphor screen. The gel image was generated using the PhosphorImager Model 400E and ImageQuant software v. 3.15 (Molecular Dynamics). The numbers to the left of each gel indicate the number of analog residues in each of the extension products.

three alleles are due to single base differences in the codons for amino acids 112 and 158 (8). Figure 4 shows detection of these polymorphisms using primer extension products approximately 500 nt in length. Synthesizing the bottom strand with an A analog or the top strand with a T analog, it is possible to distinguish all six genotypes for this three allele system. When the C analog is used in the primer extensions, however, it has not been possible to resolve distinct bands. Figure 4 shows the best separation we have been able to achieve so far with the C analog. The reason A and T work, but not C, is that the 538 bp amplified segment of *APOE* is 75% GC. For the $\epsilon 3$ allele, the A analog extension

product contains 54 biotinylated A's out of 521 nucleotides incorporated (10.3%). For the T analog the ratio is 53/477 (11.1%), and for the C analog the ratio is 162/477 (34.0%). For the C analog products, the heterozygote bands are noticeably broader than for the homozygote bands, indicating that the different species are starting to be resolved. Thus, it seems that we are obtaining full length products, but have reached the limits of resolution for this particular gel system. The analysis of the C analog products does show that it is possible to reduce electrophoresis time even further by lowering the ionic strength of the running buffer.

DISCUSSION

This mobility shift assay requires enzymatic preparation of a primer extension product which is: a) precisely full length; and b) fully substituted at every expected site with the mobility-shifting nucleotide analog. The success of the assay therefore depends critically on the mobility-shifting analog and on how efficiently it is incorporated by the DNA polymerase being used. If the mobility-shifting analog is incorporated too slowly, misincorporation of the other three unmodified dNTPs will produce a mixture of partially substituted primer extension products. If a site on the template strand requires adjacent incorporation of too many mobility-shifting analogs, the polymerase may not be able to extend through this site. If the mobility-shifting analog contains impurities which are incorporated, the primer extension product probably will not be electrophoretically homogeneous. An early version of this assay (1) used commercially available biotin-11-dUTP and *Taq* DNA polymerase. For unknown reasons, commercially available biotin-11-dCTP gave much poorer results than biotin-11-dUTP. An alkynyl analog of biotin-11-dCTP was prepared and found to give superior results. Using these biotinylated nucleotides with 11-atom linker arms, it was difficult to resolve single base differences in primer extension products longer than 200 nt. Therefore, the use of other analogs was explored in order to increase the utility of the mobility shift assay.

Nucleotide analogs which have large groups linked through the 5-position of pyrimidines or the 7-position of 7-deazapurines can be surprisingly good substrates for certain DNA polymerases (11;17–19). A versatile synthetic route to a family of such nucleotides has been developed (2,20). These nucleotide analogs have a flexible hydrophilic polyethylene glycol linker attached to the base via an acetylenic bridge. A variety of groups have been attached to the terminal amino group of these nucleotides for testing in the mobility shift assay. In general, most of these nucleotide analogs were incorporated into full length primer extension products which appeared to be homogeneous. Although these properties fulfill the above criteria for a successful mobility-shift assay, the modified primer extension products unfortunately gave significantly broader bands than normal nucleic acids upon electrophoretic analysis. Although these broader bands could be due to the presence of minor impurities or other problems, we assume that the broader bands are an intrinsic property of these highly modified nucleic acids. Whatever the reason, in practice, nucleotides derivatized with biotin gave products with tighter bands than those with fluorescent dyes (21), a simple acetyl group, or no group on the amine at all. The relationship between band tightness and the structure of the mobility-shifting group is currently unclear, but the following structural properties might be expected to favor tighter bands: a) minimal hydrophobicity, so that multiple conformers cannot be stabilized by hydrophobic interactions; b) some rigidity, to minimize the number of conformations the mobility-shifting group can assume. (Rigidity too near the nucleoside, however, could prevent the mobility-shifting group from adopting a conformation which is acceptable to the polymerase.) This analysis suggests that biotin itself has no unique features and that superior mobility-shifting groups might be discovered.

Detection of single base differences requires electrophoretic resolution of primer extension products containing N and N+1 mobility-shifting nucleotide analogs. In the absence of increased band broadening, a larger mobility-shifting group should increase

the separation of bands and permit resolution of larger primer extension products. In an attempt to maximize mobility shifting and/or define an upper size limit for acceptance of mobility-shifting groups by DNA polymerases, biotin with a 14-atom linker was attached to alkynylamino nucleotides with a 22-atom linker derived from hexaethylene glycol (1a, 1c, and 2a in Fig. 1). The resulting biotin-36 dNTPs (1b, 1d, and 2b in Fig. 1) were incorporated by *Taq* DNA polymerase, but a significant fraction of the primer extension products were not full length. When *Tli* DNA polymerase was tried, however, predominantly full length products were observed. These biotin-36 analogs have the largest mobility-shifting group (MW 771) tried so far. The acceptance of the biotin-36 analogs by *Tli* DNA polymerase is especially illustrated by the following examples. In the analysis of the *APOE* segment, up to 163 biotin-36-dC residues are incorporated with very little premature termination observed. The analysis of the GC-rich *CYP2D6* segment required that 5 biotin-36-dC residues be incorporated in a row. In the analysis of a 288-nt AT-rich segment, 7 biotin-36-dU residues were incorporated in a row with little apparent stalling (data not shown). This ability to generate full length products is crucial to the use of the mobility shift technique.

The 36-atom linker arm causes a much greater shift per residue making it possible to detect single base differences in products of 300–500 nt. However, this greater electrophoretic retardation also causes a problem. DNA strands containing a substantial number of biotin-36 residues are electrophoresing as such high molecular weight species that we are reaching the limits of resolution for a conventional sequencing gel. Thus, in the 300–400 nt range, it should be possible to analyze most DNA segments. In the 400–500 nt range, however, the ability to resolve distinct bands on a gel will depend on nucleotide composition. Different electrophoresis conditions, different gel matrices, or different types of gels, such as capillary gels or the ultrathin gel system described by Brumley and Smith (22), may be able to resolve these higher molecular weight species and thus extend the capability of the mobility shift technique even further. With the present gel system, the results shown here indicate that reducing the thickness of the gel and reducing the ionic strength of the running buffer are relatively simple measures that can be used to minimize electrophoresis time.

The use of *Tli* DNA polymerase eliminated shadow bands due to the addition of non-templated nucleotides, but shadow bands are still sometimes observed. For example, the C analog products in Figure 3 show faint shadow bands migrating just above and below the main bands. There are a number of possible causes for the appearance of shadow bands. First, the nucleotide analogs used in this procedure must be of very high purity. Even a small amount of contamination with unsubstituted nucleotide will lead to fragments migrating faster than the major band because a mobility-shifting analog is not at every available position. These extraneous bands will always migrate faster than the main band and should be observed regardless of the DNA segment being analyzed. Second, nucleotide misincorporation can lead to shadow bands that run slower or faster than the main band. This misincorporation could occur either during the amplification process or during the labeled primer extension reaction. Because misincorporation rate probably depends on sequence context, this type of shadow band will be DNA segment and analog specific. This may be the reason for the shadow bands in the C analog products of Figure 3. Finally, because this assay uses PCR-amplified templates, there is always the possibility that the starting

genomic DNA or some other component is contaminated with a different allele.

The chief advantage of the mobility shift assay is that it can detect any single nucleotide change using conventional primer extension reactions and sequencing gels. Using just biotin-36-dCTP and -dUTP, any single nucleotide substitution can be detected by analyzing both strands. By also using biotin-36-dATP, only one strand needs to be analyzed to detect any substitution because any change in the number of G's must also affect the number of one of the other three nucleotides. By analyzing both strands with three analogs, it should be possible to analyze DNA segments as large as 1000 bp. This can be done by using outside primers to amplify a segment of about 1000 bp and then using central primers on opposite strands to synthesize primer extension products of about 500 nt. The analysis of Tay Sachs mutations in Figure 2 illustrates two additional points about the mobility shift assay. First, the use of fluorescence-labeled primers is one way to achieve nonradioactive detection. Another way would be to transfer the DNA strands from the gel to a nylon or nitrocellulose support and then use one of the many detection systems available for visualizing biotinylated DNA. Second, by using primer extension products that migrate at different points in the gel, mutations in multiple DNA segments can be assayed in a single sample. The use of differently colored fluorescent primers would be another way to multiplex the mobility shift analysis.

Another hallmark of the mobility shift assay is its ability to detect previously unknown mutations and polymorphisms. Other techniques capable of detecting novel mutations include RNase cleavage (23) or chemical cleavage (24) at mismatches, denaturing gradient gel electrophoresis (25, 26), single-stranded conformation polymorphism (SSCP) analysis (27), and heteroduplex analysis (28). The use of mobility-shifting analogs has two advantages over these other techniques. First, these techniques do not always detect every nucleotide difference. For example, White *et al.* (28) have reported a case where heteroduplex analysis detected eight out of nine point mutations, but SSCP detected only one out of the nine. In principle, mobility-shifting nucleotide analogs can be used to detect all possible single nucleotide differences. Whether this proves true in practice awaits the analysis of a larger number of DNA segments. The second reason for preferring mobility-shifting analogs is that the procedure is easily adapted to the analysis of different DNA segments. The RNase and chemical cleavage methods require optimization of reaction conditions for the particular DNA fragment being analyzed. Denaturing gradient gel electrophoresis and SSCP analysis require optimization of gel conditions. The use of a standard primer extension reaction and the use of standard sequencing gels make it more likely that the mobility-shifting analog procedure can be easily applied to a wide variety of fragments. This adaptability is particularly useful when searching for polymorphisms to be used in genetic mapping or screening a large number of different exons for mutations.

An alternative to using mobility-shifting analogs or the above competing techniques is direct sequencing of PCR products. Sequencing and analysis with mobility-shifting analogs both use primer extension reactions and denaturing gels, but the mobility shift method is less laborious in both sample processing and data analysis. The ability to multiplex the mobility shift procedure (as in Figure 2) can be used to increase efficiency even more relative to sequencing. Another problem with sequencing is the variability in reactions and in data interpretation. PCR fragments

are notorious for their intractability as sequencing templates. Thus, sequencing reaction conditions must be optimized for each particular DNA segment being analyzed. Ambiguities in data interpretation can also cause problems. For example, the appearance of bands at the same position in two lanes can be a sequencing artifact or caused by heterozygosity at that position. In fact, one of the strong points of the mobility-shifting analog method is that it gives a clear and unambiguous indication of heterozygotes, namely, the appearance of two bands of equal intensity.

ACKNOWLEDGEMENTS

We are grateful to Dr. Robert Desnick of the Mount Sinai Medical Center for providing genomic DNA from individuals carrying the Tay Sachs mutations, to Dr. Richard Spielman of the University of Pennsylvania for providing the CEPH family genomic DNA, and to Dr. A.J.Cocuzza for the gift of intermediates used in the synthesis of biotin-36-dATP.

REFERENCES

- Kornher, J. S. and Livak, K. J. (1989) *Nucleic Acids Res.* **17**, 7779–7784.
- Hobbs, F. W., Robinson, D. and Yarem, J. A. (manuscript in preparation).
- Myerowitz, R. and Costigan, E. C. (1988) *J. Bio. Chem.* **263**, 18587–18589.
- Myerowitz, R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3955–3959.
- Navon, R. and Proia, R. L. (1989) *Science* **243**, 1471–1474.
- Myerowitz, R., Piekarz, R., Neufeld, E. F., Shows, T. B. and Suzuki, K. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7830–7834.
- Kimura, S., Umeno, M., Skoda, R. C., Meyer, U. A. and Gonzalez, F. J. (1989) *Am. J. Hum. Genet.* **45**, 889–904.
- Heim, M. and Meyer, U. A. (1990) *Lancet* **336**, 529–532.
- Emi, M., Wu, L. L., Robertson, M. A., Myers, R. L., Hegele, R. A., Williams, R. R., White, R. and Lalouel, J. (1988) *Genomics* **3**, 373–379.
- Paik, Y., Chang, D. J., Reardon, C. A., Davies, G. E., Mahley, R. W. and Taylor, J. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3445–3449 (1985).
- Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., Cocuzza, A. J., Jensen, M. A. and Baumeister, K. (1987) *Science* **238**, 336–341.
- Hobbs, F. W. (1991) U. S. Patent **4,997,928**.
- Garoff, H. and Ansoerge, W. (1981) *Anal. Biochem.* **115**, 450–457.
- Clark, J. M. (1988) *Nucleic Acids Res.* **16**, 9677–9686.
- Kagimoto, M., Heim, M., Kagimoto, K., Zeugin, T. and Meyer, U. A. (1990) *J. Biol. Chem.* **265**, 17209–17214.
- Davignon, J., Gregg, R. E. and Sing, C. F. (1988) *Arteriosclerosis* **8**, 1–21.
- Langer, P. R., Waldrop, A. A. and Ward, D. C. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6633–6637.
- Ward, D. C., Langer, P. R. and Waldrop, A. A. (1987) U. S. Patent **4,711,955**.
- Schmitz, G. G., Walter, T., Seibl, R. and Kessler, C. (1991) *Anal. Biochem.* **192**, 222–231.
- Hobbs, F. W. (1989) *J. Org. Chem.*, **54**, 3420–3422.
- Chelsky, D., Begy, C., Talmadge, C., Trainor, G. and Hobbs, F. (manuscript in preparation).
- Brumley, R. L. and Smith, L. M. (1991) *Nucleic Acids Res.* **19**, 4121–4126.
- Myers, R. M., Larin, Z. and Maniatis, T. (1985) *Science* **230**, 1242–1246.
- Cotton, R. G. H., Rodrigues, N. R. and Campbell, R. D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4397–4401.
- Myers, R. M., Lumelsky, N., Lerman, L. S. and Maniatis, T. (1985) *Nature* **313**, 495–498.
- Sheffield, V., Cox, D. R., Lerman, L. S. and Myers, R. M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 232–236.
- Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. (1989) *Genomics* **5**, 874–879.
- White, M. B., Carvalho, M., Derse, D., O'Brien, S. J. and Dean, M. (1992) *Genomics* **12**, 301–306.