Automated Detection of Point Mutations by Electrophoresis in Peptide-Nucleic Acid-Containing Gels

Research Report

BioTechniques 27:798-808 (October 1999)

Gabor L. Igloi
Universität Freiburg,
Freiburg, Germany

INTRODUCTION

Advances in molecular diagnostics, forensic science and gene therapy require a routinely applicable and preferably automated sifting of a large number of DNA samples for possible single nucleotide polymorphisms (SNP) or point mutations. The available techniques, such as single-strand conformational polymorphism (SSCP), ligase chain reaction or oligonucleotide ligation assay rely largely on the analysis of appropriate products from the polymerase chain reaction (PCR), and most involve extensive post-amplification manipulations or interpretive skills that are not compatible with automation. Sequence-specific hybridization of nucleic acids (16) forms an integral part of several procedures for detecting sequence variations. The conventional approach, involving membrane-bound targets that interact with probes in solution, cannot meet the demands of the large-scale, rapid screening dictated by many of today’s applications.

Peptide nucleic acids (PNAs) are synthetic, electronically neutral chimeras of nucleobases linked to a polyamide backbone (4). This spacing permits the bases to form standard base pairs with natural nucleic acids. The discriminatory qualities of PNA/DNA interactions in detecting mutations have been previously used in the form of PCR clamps (10,11). Carlsson et al. (2), Demidov et al. (3) and Rose (14) have shown that DNA/PNA duplexes can be separated from unbound DNA by electrophoresis, while Perry-O’Keefe et al. (12) have proposed a pre-gel hybridization system based on the same observation.

The lack of the phosphodiester linkage, leading to an electronically neutral species, has important consequences for the base-pairing potential of PNA. The greater stability of PNA/nucleic acid duplexes, together with the salt independence of hybridization that has been used for SNP detection in MALDI-TOF analyses (15), has been ascribed to the lack of backbone charge repulsion. These two factors, electronic neutrality and salt independent base pair specificity, are the properties of PNA that also make them ideal candidates as ligands in an affinity electrophoretic partnership (6).

Entrapment of PNA into native polyacrylamide gels creates an affinity electrophoretic matrix that selectively retards complementary nucleic acids in a sequence-specific manner. It is highly sensitive to single base mismatches (8) and can be readily automated using commercial fluorescence-based DNA sequencers. By defining the essential parameters of the interaction in a model system, the concept of hybridization in real time has been applied to clinically relevant hetero- and homozygous mutation detection.

MATERIALS AND METHODS

PNAs were obtained from PerSeptive Biosystems (Framingham, MA, USA) as crude products. They were synthesized with one or more N-terminal [2-(N-boc-2-aminoethoxy)ethoxymethyl] acid linkers or spacers, as recommended by the manufacturer, to prevent intramolecular rearrangement.
They were further purified by passage through a NAP5 (Amersham Pharmacia Biotech, Freiburg, Germany) column. Fractions (0.2 mL water) were collected and UV260 absorbing material combined and stored frozen. Affinity gels were constructed by mixing 9 mL 2× MDE™ gel solution (FMC Bioproducts, Rockland, ME, USA), 3.6 mL 10× TBE (1 M Tris, 0.83 M boric acid, 10 mM EDTA) together with the appropriate amount of PNA in a total volume of 36 mL. Polymerization was initiated by the addition of 130 µL 10% ammonium persulfate and 30 µL (TEMED). The gel solution was poured immediately between the plates of an ALF™ DNA Sequencer (Amersham Pharmacia Biotech) using 0.35-mm spacers. The temperature of the gel was controlled externally by connecting a Multi Temp II Thermostat Bath (Amersham Pharmacia Biotech) to the ALF thermostat plate. Gels were run routinely in 0.8× TBE at 34 W constant power.

PCR samples were obtained from human wild-type and cystic fibrosis ΔF508 mutant (homo- and heterozygous) or hemochromatosis C282Y (homo- and heterozygous) DNA cycling at 94°, 55° and 72°C for 30 s, for each of 30 cycles (GeneAmp® PCR System 9600; PE Biosystems, Foster City, CA, USA). Primers were synthesized in the auto-purification mode of a Model 3948 System (PE Biosystems). The 5′ fluoroscein versions of these primers were obtained using FluorePrime™ phosphoramidite (Amersham Pharmacia Biotech) during the synthesis on a Model 394 machine (PE Biosystems). Where necessary, a BsaI recognition site was introduced by adding the sequence 5′ TTTGTGCAG to the 5′ end of the primer. Having established the appearance of a PCR product of the expected size by conventional agarose gel electrophoresis, aliquots from the PCR were diluted 1:20 with water. For affinity electrophoresis, 3 µL of the diluted PCR product (approximately 150 fmol) were mixed with 3 µL 1% Dextran Blue (Amersham Pharmacia Biotech) in formamide, denatured at 95°C for 2 min, chilled in ice and 3 µL applied to a gel. The fluorescent signals were automatically detected by an ALF automated sequencer and could be analyzed further using the Fragment Manager software (Amersham Pharmacia Biotech). Quantitative data was expressed graphically using Microsoft® Excel® (Microsoft, Redmond, WA, USA).

RESULTS

The Cystic Fibrosis ΔF508 Deletion as a Model Point Mutation

Relationship of the PNA sequence to the PCR product. To identify the species involved in any relative shift of the observed signals in a model system, fragments differing in length, as well as in the presence or absence of a mismatch, were chosen. For this approach, a PNA 11-mer specific for the ΔF508 mutation in the human cystic-fibrosis transmembrane protein gene (13) was polymerized into a non-denaturing polyacrylamide gel.

The PNA sequence (Figure 1) formed a fully base-paired duplex, solely with the mutant complementary strand. However, despite the deletion of a codon in the mutant, the selected PNA sequence was such that on binding the wild-type complementary strand, only a single mismatch (A:A) at position 4 from the 5′ terminus could form. Thus, the model system mimicked a point mutation while permitting a clear assignment of the electrophoretic pattern to the PCR products on the basis of the 3-bp deletion (wild-type → mutant). Fractionation of PCR products. PCRs were performed for each DNA sample using a fluorescent primer that corresponded to the strand complementary to the immobilized PNA. The signal then observed corresponded to the strand labeled by its primer. The denatured products of amplification of genomic human DNA from wild-type and mutant sources generated a migration pattern in a standard non-denaturing gel (Figure 2A), in accordance with expectations: a single product from wild-type (98 bp) and homozygous mutant DNA (95 bp) and a doublet differing in 3 bases in the heterozygous form. By including PNA into a non-denaturing polyacrylamide gel and applying the same PCR samples, a specific retardation effect by the PNA for the complementary mutant strand would, according to studies on oligonucleotides (8), be expected. However, the behavior of the PCR fragments in the gel is evidently more complex. The temperature dependence of the migration of the 98/95-mer products (Figure 2B) indicated that although there had been a relative shift in signals, reducing the heterozygous doublet (cf. Figure 2A) to a singlet, all three products migrated at temperatures below 60°C to approximately the same position. This shift prevented a simple differentiation between a homozygous mutant and the wild-type species. Above 60°C, interactions with the entrapped PNA became destabilized, and the familiar pattern seen in the absence of PNA was regained (Figure 2C). A number of variables, not necessarily independent but affecting the interaction, must be considered.

The length of the DNA could be reduced successively by using the native restriction sites within the amplified fragment or those sites introduced by appropriate primer design. Eliminating 29 bases from the 3′ end with Tru9I facilitates the identification of the mutant strand by allowing an improved differentiation of the signals (Figure 2D).

The above treatment, while suggesting a structural factor that influences retardation, cannot by itself lead to a definitive conclusion. A consequence of a stepwise reduction in the PCR

Figure 1. Structural relationship between the PCR products from wild-type and mutant cystic fibrosis DNA, with respect to the PNA sequence. Spacers at the N terminus of the PNA are denoted by O, and the codon deleted in the mutant is in bold type. Note the single A:a mismatch.
product length is a gradual shift of the PNA binding region towards the DNA 3′ terminus. The observed effect might then not simply depend on the molecular size of the migrating species, but on the position of the sequence complementary to the PNA with respect to the termini. Stepwise lengthening of the 5′ end of the lower strand, by overlapping PCR primers (Figure 3A) while retaining the cleavage at the 3′ end and quantifying the resolution, leads to the conclusion that the main factor governing the interaction of the DNA with the matrix-entrapped PNA is the relative

![Figure 2](image-url)
position of the PNA binding sequence to the DNA terminus, and not the absolute length of the DNA strand. The relationship between the degree of retardation is best described empirically by a linear logarithmic function (Figure 3B). Clearly, the closer the PNA binding site approaches either terminus, the better the resolution, while the absolute fragment length shows little correlation (not shown).

For general applicability however, the design of PCR primers around a mutation of interest should not be limited by having to include appropriate natural restriction sites. Nevertheless, it may be desirable to relocate the PNA binding site towards the DNA terminus by restriction digestion, for instance. A BsgI site [5'GTGCAG(N16/N14)] is ideally suited for this purpose, as it requires only an addition of the six bases of the recognition site, plus three further 5' terminal bases to any PCR primer. Thus, the amplification, followed by BsgI treatment to remove much of the primer sequence, generates a mutant strand that is even more strongly retarded in the PNA gel than the Tru9I product (Figure 2E).

The Hemochromatosis C282Y Point Mutation

Conclusions reached from the model system above could now be verified for a clinically relevant point mutation, while wild-type and mutant strands could not be distinguished by length differences. The C282Y hemochromatosis mutation is caused by a G/A transition (5). A PNA 11-mer, fully complementary to the wild-type, was chosen (Figure 4A) to confirm the utility of the method under conditions where the mutant mismatch (G:T) had previously been shown to be the least favorable in affinity electrophoretic discrimination for oligonucleotides (8). The PNA also represents an example of a sequence with 73% GC content, whose stable hybridization to DNA [with correspondingly high melting temperature ($T_m$)] might also be expected to hinder an effective fractionation. Figure 4B demonstrates a clear and unambiguous resolution of the wild-type (retarded) from the mutant species, in contrast to standard electrophoresis in a non-denaturing gel, which was unable to distinguish between the wild-type and the mutant product (Figure 4C).

Here, the fractionation of PCR products having the PNA binding site 22 bases from the 3' terminus and complementary to the wild-type sequence, provided adequate resolution between wild-type and mutant (since no length difference had to be overcome). Again confirming the results of the cystic fibrosis model, the absolute length of the PCR product had little influence on the resolution, as was demonstrated by successfully fractionating a 451-bp product in this system (see Figure 4D compared with the control, Figure 4E).

As expected from the theory of base-pair formation and analogous to conventional membrane-based hybridization, the degree of retardation (or binding) is also influenced by a number of physical parameters that must be optimized for each system. At low temperatures, for instance, the PCR product with the single base mismatch (or other regions of sequence similarity in the PCR product) will become retarded as a result of PNA/DNA duplex formation; at high temperatures, the fully complementary species will become destabi-
lized with a decreased retardation. Moreover, the Tm of the hybrid, as is the case for DNA/DNA duplexes (1), is concentration-dependent so that the optimal resolution at a given temperature is also related to the concentration of the entrapped PNA. Figure 5 shows a graphical representation of the mobility shifts achieved at various concentrations of entrapped PNA with respect to the gel temperature. At high concentrations of PNA (200 nM), high temperatures (>60°C) are required before the nonspecific binding is overcome. At low PNA concentrations (2 nM), the same temperature leads to dissociation of specific binding and a loss of resolution. By varying the PNA concentration and the temperature, one may optimize the resolution of the observed signals and compensate for the GC content of the PNA, for instance.

**DISCUSSION**

The essentially uncharged nature of PNA permits the physical entrapment of an 11-mer that remains sufficiently immobile in an electric field to enable its interaction with complementary sequences to be studied. Using PCR primers labeled with fluorescein to generate fluorescent PCR products, one can detect the affinity electrophoretic retardation of interacting species with the aid of a conventional automated DNA sequencer. The PNA 11-mer, having the sequence of the common ΔF508 mutation in the cystic fibrosis transmembrane gene product, binds and retards only the PCR product from mutant samples. Significantly, the sequence of the PNA varied from the wild-type sequence solely by a single-base mismatch. This effectively provided a system that acts as a model for the detection of point mutations by enabling a comparison of the mobility of signals with an internal reference—the longer wild-type sequence. Being a genuine point mutation, the hemochromatosis C282Y mutation is an example where size differentiation is not possible and affinity electrophoretic mutation detection can play a significant role in diagnostic screening.

In contrast to the conventional physical parameters that determine the interaction, it has now been found that effective base pairing has rather stringent structural requirements. While the overall length of the PCR product has little influence on PNA/DNA duplex formation, the location of the PNA binding region (with respect to either terminus) is important. Although not rigorously defined, there appears to be an inverse logarithmic relationship between the

![Figure 4. Section of the human HLA-H DNA sequence encompassing the hemochromatosis mutation.](image)
strength of the interaction and the distance of the binding region from the DNA ends. One can envisage at least two models to explain such structural effects. It is possible that the PNA is entrapped in the polyacrylamide pores in such a way that only the ends of the migrating DNA strands can make effective contact by penetrating into the matrix. Alternatively, following denaturing, the ends of the DNA are more likely to remain in the single-stranded form required for PNA/DNA duplex formation than more central regions.

Whatever the mechanism of the interaction, the structural constraints demanded by the affinity electrophoretic system can be readily met by an appropriate primer design. Clearly, shorter PCR products are preferable because they require shorter electrophoretic run times. Positioning the PNA binding region near to the 5′ or 3′ DNA terminus is achieved by selecting PCR primers so the unlabeled primer lies almost adjacent to the PNA region, while the labeled primer is designed to be compatible within the desired length constraints. If the resolution is inadequate, the strategy of incorporating a BssGI site into the unlabeled primer, as described, provides a universal tool for relocating the PNA binding region towards the 3′ terminus and enhances the fractionation achieved.

The PNA sequence is designed to incorporate the point mutation. The nature of the retardation corresponds to the selected sequence and may be either the wild-type or the mutant; in either case, only the complementary, fully matched strand is bound and retarded by the matrix. In the case of the heterozygous sample, only the labeled complementary strand is shifted on binding to the PNA, while the mismatched strand is not retarded. In view of results obtained with oligonucleotides (8), the mismatch has, in the examples described here, been placed at the most discriminating position, base number four, in the PNA sequence. However, it is likely that other positions would discriminate adequately. One should also note the difference in the G+C content of the two PNAs used. The cystic fibrosis model could be resolved with 36% G+C, while the hemochromatosis mutation encompassed a 73% G+C region. The versatility of the analysis permits such extremes by allowing the physical parameters (PNA concentration and temperature) to be adjusted accordingly and characteristically for each PNA/DNA interaction.

The fluorescent label required for detection may be introduced into the PCR product using primers labeled 5′ terminally during synthesis or enzymatically at the 3′ terminus (7,9). Alter-

![Figure 5. Temperature dependence of the migration of the retarded species, with respect to the PNA concentration. The time point of the appearance of the retarded signal (wild-type) with respect to the unretarded signal (C282Y mutant) was determined at various temperatures for entrapped PNA concentrations of 2 nM (●), 20 nM (×) or 200 nM (○).]
natively, the primer that will form the labeled strand may be extended 5’ terminally with a universal sequence, e.g., the T7 promoter sequence, and the label introduced by PCR using the corresponding fluorescent universal primer.

The use of fluorescent primers is important for detecting hybridization in real time using an automated DNA sequencer. The highly sensitive system used in this study, requiring low concentrations of PNA and the application of only a small fraction of the PCR product obtained from genomic DNA, can be used to analyze many samples simultaneously. Indeed, it is conceivable that by including a mixture of various PNA sequences in the gel and using PCR primers labeled with different fluorophores (to give products with labels that correspond to one of the corresponding fluorescent dyes), a multiplexing system can be developed to detect several mutations simultaneously in single or multiple genes. The amount of information available from a single gel run is then only limited by the number of detectable fluorophores. Other commercial fluorescence-based sequencers may be used analogously, provided the raw data is available for inspection or the software is not limited to sequence analysis. The potential for a corresponding fractionation by capillary electrophoresis has been established (unpublished results).

Applying an automated DNA sequencer for data acquisition is the first step in automating large-scale mutation screening and, in general, permits the visualization of hybridization processes in real time.

ACKNOWLEDGMENT

This investigation would not have been possible without the assistance of Mrs. E. Schieffermayr. I am grateful to Dr. G. Scherer for valuable discussions and to Dr. E. Natt for providing the genomic DNA samples. Financial support was, in part, by the Deutsche Forschungsgemeinschaft (SFB 388).

REFERENCES


Received 6 April 1999; accepted 11 June 1999.

Address correspondence to:
Dr. Gabor L. Igloi
Institut für Biologie III
Universität Freiburg
Schänzlestr. 1
D-79104 Freiburg, Germany
Internet: igloi@oligo.biologie.uni-freiburg.de