Application of natural and amplification created restriction sites for the diagnosis of PKU mutations

H.G.Eiken*, E.Odland, H.Boman, L.Skjelkvåle[1], L.F.Engebretsen and J.Apold
Department of Medical Genetics, Haukeland Hospital, University of Bergen, N-5021 Bergen and
[1]Department of Pediatric Research, University of Oslo, Oslo, Norway

Received January 22, 1991; Revised and Accepted March 1, 1991

ABSTRACT

PCR amplification, either conventional, or as site directed mutagenesis using primers with mismatched 3'-ends, followed by restriction endonuclease digestion, provides rapid, non-isotope assays of known mutations in the human phenylalanine hydroxylase gene. Such assays were shown to have the potential to detect all of the 18 presently reported phenylketonuria mutations. The practical applicability of this approach was demonstrated for eight mutations in Norwegian phenylketonuria patients, among them the most common ones.

INTRODUCTION

In combination with PCR amplification, restriction enzyme analysis forms a rapid and general technique to identify gene lesions that change a restriction enzyme recognition site. In most cases, however, there is no restriction enzyme available that directly recognizes the mutation. In such instances appropriate restriction sites can be artificially created by site directed mutagenesis in the amplification step (1), referred to as amplification created restriction sites (ACRS).

In general, the combined PCR/restriction enzyme method is a two step process. Firstly, a set of specific primers is annealed to the target DNA to give a specific PCR product, i.e. one single band in analytic gel electrophoresis. Secondly, after the PCR product has been digested with a known restriction enzyme and subjected to electrophoresis, a specific pattern of bands appears on the gel. These diagnostic criteria have so far been applied to a number of known mutations in single copy genes (2,3,4). However, the creation of an artificial restriction site, either in the wild-type or in the mutated gene, has only been performed in a few cases: e.g. the ΔF508 mutation of the cystic fibrosis gene (5), point mutations of codon 12 of the Kirsten ras oncogene (1), and in polymorphisms of the alcohol dehydrogenase loci (7).

Here we show how natural, or, if necessary, artificial restriction sites can be established in known DNA sequences. This approach provided a theoretical basis for diagnosing all published mutations in the phenylalanine hydroxylase (PAH) gene. These are, with few exceptions, single base mutations, causing phenylketonuria (PKU), the most common inborn error of amino acid metabolism. The practical feasibility of this approach was then demonstrated for eight selected mutations, comprising the most common ones in Norwegian PKU patients.

MATERIALS AND METHODS

Computer software and analysis

Searches for restriction sites discriminating between morbid and normal PAH genes were performed by the computer program DNASIS (Pharmacia-LKB). The DNA sequences adjacent to the PKU mutation were examined for naturally occurring restriction enzyme recognition sequences. This search was carried out with the recognition sequences of a selection of 45 hexanucleotide, 6 pentanucleotide, and 11 tetranucleotide, commercially available enzymes, excluding isoschizomers. For these three groups of enzymes, the search included five, four, and three nucleotides to each side of the mutation site, respectively.

If no natural recognition sequence was found, each of the ten, eight, and six nucleotides in the frames encompassing the PKU mutation, respectively, were individually changed (artificially 'mutated'), and a permutated set of searches for new recognition sequences performed. For each new mutation search using this software, a total of only 80 characters had to be entered. When more than one enzyme was found, we chose, for the practical experiment, the enzyme that we found most favorable with respect to expenses, availability, primer design, and the length of the ACRS products.

PCR amplification and restriction enzyme analysis

All exons were amplified in 100 μl reaction volumes containing < 0.5 μg genomic DNA, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatine, 0.12 mM of each dNTP, 0.6 μM of each primer, and 2.5 units of Taq polymerase (Perkin-Elmer Cetus Co.). A total of 30 cycles were carried out at 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, followed by incubation at 72°C for 7 min. All primers were purchased from Genesys, MedProbe Inc. The primer sequences were complementary to the flanking sequences of exons 1, 5, 7, 11, and 12 (8, 12, 10, and 11). The two primers used in ACRS had

* To whom correspondence should be addressed

been modified to (5A-MspI primer) 5'-ATCCTGTGTA-CCGTGCAAGC-3' and (12B-RsaI primer) 5'-CTCGTAA-GGTGTAAATTACGTA-3'(see Figure 1).

Samples of 40 μl of the PCR reactions were subjected to digestion with the discriminating restriction enzyme in the appropriate buffer for 2–4 hours. Any doubt about the activity of the restriction enzyme was resolved by addition of lambda phage DNA. One third of the digested sample, and an equal aliquot of the undigested PCR-product, were analyzed in 3% Nusieve agarose (FMC Co.).

Patients

DNA isolated from peripheral blood samples of 49 Norwegian patients with classical PKU, was analyzed, including 43 patients with PAH DNA haplotypes described previously (9).

RESULTS

There are more than 100 different restriction enzyme recognition sequences known. Approximately 1/6 of the basepairs in the human haploid genome is covered by a naturally occurring restriction enzyme recognition site. Less than 20% of the mutations would then be directly detectable by an available restriction enzyme. However, by mutating, in the PCR process, one basepair near the site of the PKU mutation, the possibility to cover the PKU mutation site by a restriction site was greatly increased.

The theoretical probability of success (P) for x unique restriction enzyme sequences is

$$P=1-[(N-N')/N]^x \qquad (i)$$

where N is the total number of different sequence combinations of n nucleotides that a specific n-basepair recognizing type II enzyme will discern. Obviously, N equals 4^n.

N' is the number of possible permutated nucleotide combinations given in n different frames stretching (n−1) nucleotides to each side of the mutation. These (n−1) nucleotides may be any of the four bases. The number of possible nucleotides at the site of the mutation, however, is restricted to two: the given mutated base or its corresponding wild-type equivalent. For a restriction enzyme recognizing n nucleotides, the number of possible permutations (N') is $2n[3(n-1)+1]$, including the naturally occurring sequences.

As an example, consider the probability of success for finding a natural, or creating an artificial, restriction enzyme recognition site encompassing a given PKU single base mutation. In this example, we will use 11 tetranucleotide recognizing enzymes (x=11, n=4). For an enzyme recognizing four basepairs, there are a total of 256 different combinations of possible recognition sites (N=256). By employing the considerations above, the number of recognition sites found naturally, or created by a single artificial mutation, in n=4 different frames encompassing the morbid mutation and extending maximally 3 bases (n-1) to each side, is 80 (N'=80). Thus, from (i) it follows that the theoretical probability of success in this case is 98.4%.

The corresponding probabilities of success for six pentanucleotide and 45 hexanucleotide recognition enzymes are 55.7% and 88.5%, respectively. These probabilities are not independent and should be considered with caution. However, a search performed with a combination of these enzymes, is likely to yield one or more possible solutions in nearly all cases.

The results of a computer search using the given selection of enzymes are summarized in Table 1. On the basis of the results

Table 1. List of restriction enzymes that discriminate between normal and all reported morbid PAH genes. Further information on the PKU mutations are given in ref. (6). E: Exon. I: Intron.

NATURALLY OCCURRING TARGET SEQUENCES:

Mutation	Pos. in cDNA	E/I	Population	Reference	Restriction enzyme
M1V	1	E1	Canada	John et al. 1989	NlaIII[a]
F39L	339	E2	Australia	Forrest et al. 1989	MboIIB
R252W	976	E7	Mediterr.	Abadie et al. 1989	AvaI
R261Q	1004	E7	Europe	Okano et al. 1990	HinfI
G272ter	1036	E7	Norway	Apold et al. 1990	BamHI[b]
			Sweden	Svensson et al. 1990	"
364del	1312–14	E11	Sweden	Svensson et al. 1990	HindIII[c]
R408W	1444	E12	Europe	DiLella et al. 1987	MnlI[d]

TARGET SEQUENCES CREATED BY MUTAGENESIS:

Mutation	Pos. in cDNA	E/I	Population	Reference	Restriction enzyme
R111ter	553	E3	Orientals	Wang et al. 1989	MaeI
R158Q	695	E5	Europe	Okano et al.1990	MspI
Y204C	833	E6	China	Wang et al. 1989	MaeI
R243ter	949	E7	Europe	Okano et al. 1989	TaqI
L255S	986	E7	Am.Blacks	Hofman et al. 1987	DdeI
E280K	1060	E7	Lyonnet	Abadie et al. 1989	TaqI
P281L	1064	E7	Europe	Okano et al. 1989	MspI
F299C	1118	E8	"	Okano et al. 1989	HindIII
L311P	1154	E9	Germany	L.-Konecki et al.1988	MaeI[e]
R413P	1460	E12	China	Wang et al. 1989	TaqI
Splice12	–	I12	Europe	DiLella et al. 1986	RsaI

a) ref. (8), b) ref. (9), c) ref. (10), d) ref. (4), e) the mutation generates an MspI site (ref. (21)), and may alternatively be used for diagnosis.

of the computer search, a suitable diagnostic test for the common PKU mutations in Caucasians was envisioned. For the Splice12 (11) and R158Q (12) mutations, prevalent among European PKU patients, no natural restriction enzyme recognition site was found.

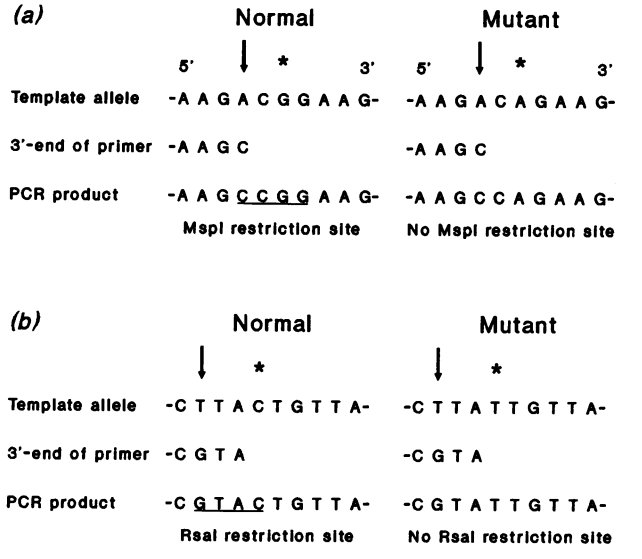

Figure 1. Strategy for the detection of the R158Q mutation (a) and the Splice12 mutation (b). The arrows indicate the location of the nucleotide mismatch, while an asterisk (*) indicates the position of the PKU mutation.

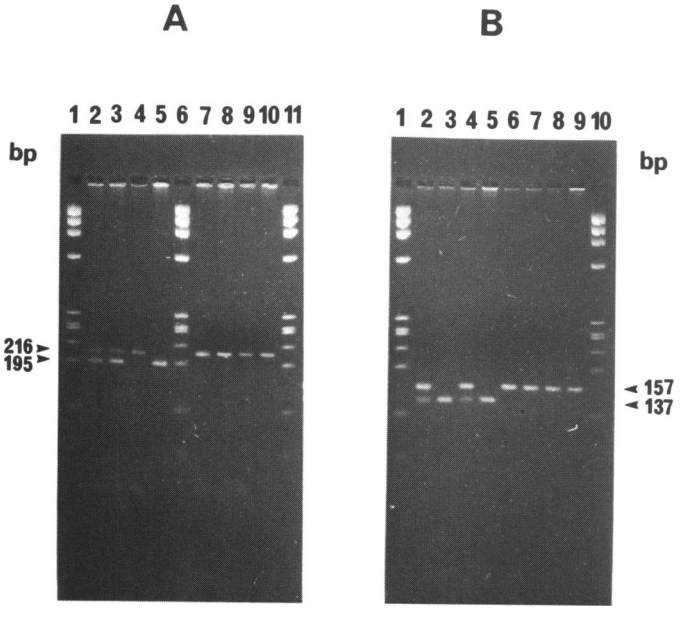

Figure 2. The ACRS approach applied to families with the Splice12 (**A**) and R158Q (**B**) PKU mutations. DNA was first amplified and digested with RsaI and MspI, respectively. Then aliquots of the samples were electrophoresed in 3% NuSieve agarose for 2 h at 7.5 V/cm. **A:** Lanes 2 and 3: The parents, both heterozygous for the Splice12 mutation. Lane 4: The PKU patient homozygous for this mutation. Lane 5: An unaffected, non-carrier sister. Lanes 7, 8, 9, and 10 show the amplification products before RsaI digestion, in the same order. Lanes 1, 6, and 11: HaeIII-digested ϕX174 DNA molecular weight marker. **B:** Lanes 2 and 4: The father and the PKU patient, heterozygous for the R158Q mutation. Lanes 3 and 5: The mother and an unaffected sister do not have the R158Q mutation. Lanes 6, 7, 8, and 9 show the amplification products before the digestion with MspI, in the same order. Lanes 1 and 10: HaeIII-digested ϕX174.

The computer search indicated that such sites could be created by ACRS.

Following this search for useful restriction enzymes, we designed diagnostic tests for all the common PKU mutations in Caucasians. Among these, tests based on restriction enzyme cleavage have been previously reported for the mutations M1V, G272ter, 364del, and R408W (8,9,10,4). For the R261Q (12) and R252W (13) natural restriction enzyme recognition sites were found (Table 1). The Splice12 and R158Q mutations are prevalent in European populations, and were chosen as examples of the experimental design for the ACRS assays. The theoretic basis for the detection of these two mutations is shown in Figure 1.

Figure 2 shows the ACRS laboratory assays for the Splice12 and the R158Q mutations. The ACRS products, differing by one basepair from genomic DNA, are visualized on an electrophoresis gel after restriction enzyme digestion. The mutant fragments remain uncleaved, whereas the normal fragments are cleaved by the appropriate restriction enzyme. The lengths of the restriction fragments were as predicted from the DNA sequences (195 + 21 bp for the Splice12 mutation, and 137 + 20 bp for the R158Q mutation), and mendelian transmittance is demonstrated. Also, these mutations were found on the haplotype background expected from the known linkage disequilibrium between the Splice12 mutation and PAH haplotype 3 (11), and the R158Q mutation and haplotype 4 (12).

The results of the assays of the eight mutations selected for such tests in 49 Norwegian PKU patients showed that five of the mutations were present in this population sample: R158Q (n=3), R261Q (n=7), G272Ter (n=14), R408Q (n=15), and Splice12 (n=18).

DISCUSSION

The use of DNA amplification, followed by restriction enzyme cleavage, forms a reliable and rapid, non-radioactive diagnostic assay of known mutations. When natural restriction sites are not available, artificial restriction sites may be created in the amplification step (ACRS). We show that this method may be nearly universally applicable. The practical usefulness of the approach is illustrated by the assay of eight different mutations in 49 patients presumably homozygous for a PKU. Two of the mutations required ACRS.

We find that ACRS is the most satisfactory diagnostic test in diseases with multiple mutations. This is due to the robustness and safety emerging from the two independent steps of the assay: Firstly, a specific amplification product is generated that can be checked for correct fragment length. Secondly, the restriction enzyme analysis of the amplification product generates DNA fragments, whose appropriate sizes can also be examined by gel electrophoresis.

Misdiagnosis may occur if there, in addition to the mutation assayed, is an unknown, silent mutation in the recognition sequence. The probability for such misdiagnosis is very small, since most target DNA is highly conserved exon sequence, and apply also to the alternative methods.

We have shown that the theoretic probability of establishing these assays for any known mutation may be close to 100%. However, due to the variable kinetics of mismatches, depending on their distances from the 3'end of the oligonucleotide used, mismatches positioned 3, 4, and 5 nucleotides, respectively, from the 3' end of the primer should give the best results (14). Problems may arise when the 3' ultimate or penultimate base

of the primer is chosen as the mismatch (15). The efficiency of amplification depends on the content of G-C in contrast to A-T pairs in the neighboring positions, and on the nature of the mismatch (C-G=T-G>G-G>A-G; see ref. (16,17)). For unfavorable constructs the assay may fail, and accordingly the probability of constructing a successful assay for a mutation may be somewhat reduced. However, recent reports suggests that the versatility of PCR-based site directed mutagenesis may be extended to include also the use of primers having a nonmatched 3′ end (16).

ACRS is a favorable alternative to allele specific oligonucleotide hybridization (ASO) (18), and the amplification refractory mutation system (ARMS) (15,19,20). ASO has been widely used, but it requires the use of isotope, as well as an extra set of oligonucleotides for the mutation detection following amplification. ARMS is a one step assay, without the safety and robustness from the additional restriction enzyme analysis in A-CRS. It depends critically on an efficient initiation of the elongation of the oligonucleotide primers that provide amplification of the wild-type allele. The ARMS method is sensitive to changes in annealing temperature and target DNA concentrations (19). By the ACRS method, but not by the ASO or the ARMS methods, also heterozygotes may be identified in a single experiment, with only one set of primers. Thus, by A-CRS, family analysis, with the demonstration of mendelian transmission of the mutations, may be performed in some few experiments.

ACKNOWLEDGEMENTS

We wish to thank professor Savio L.C.Woo for his continuing support, including the sharing, prior to their publication, of primer sequences. The skilful technical assistance of Mrs. Olaug F.Bratbak is also gratefully acknowledged. This work was in part supported by the Norwegian Research Council for the Sciences and the Humanities.

REFERENCES

1. Haliassos,A., Chomel,J.C., Tesson,L., Baucis,M., Kruh,J., Kaplan,J.C. and Kitzis,A.(1989) Nucleic Acids Res., **17**, 3606.
2. Saiki,R.K., Scharf,S., Faloona F., et al. (1985) Science, **230**, 1350–1354.
3. Gencic,S., Abuelo,D., Ambler,M. and Hudson,L.D. (1989) Am. J. Hum. Genet., **45**, 435–442.
4. Eiken,H.G., Odland,E., Hordnes,K., Boman,H. and Apold,J. (1990) Clin. Chem., **36**, 809–810.
5. Friedman,K.J., Highsmith,W.E., Prior,T.W., Perry,T.R. and Silverman,L.M. (1990) Clin. Chem., **36**, 695–696.
6. John,S.W.M., Rozen,R., Scriver,C.R., Laframboise,R. and Laberge,C. (1990) Am. J. Hum. Genet., **46**, 970–974.
7. Groppi,A., Begueret,J. and Iron,A. (1990) Clin. Chem., **36**, 1765–1768.
8. John,S.W.M., Rozen,R., Laframboise,R., Laberge,C. and Scriver,C.R. (1989) Am. J. Hum. Genet., **45**, 905–909.
9. Apold,J., Eiken,H.G., Odland,E., Fredriksen,Å., Bakken,A., Lorens,J.B. and Boman,H. (1990) Am. J. Hum. Genet., **47**, 1002–1007.
10. Svensson,E., Andersson,B. and Hagenfeldt, L. (1990) Hum. Genet. **85**, 300–304.
11. DiLella,A.G., Marvit,J., Lidsky,A.S., Güttler,F. and Woo,S.L.C. (1986) Nature, **322**, 799–803.
12. Okano,Y., Wang,T., Eisensmith,R.C., Steinmann,B., Gitzelmann,R. and Woo,S.L.C. (1990) Am. J. Hum. Genet., **46**, 18–25.
13. Abadie,V., Lyonnet,S., Maurin,N., Berthelon,M., Caillaud,C., Giraud,F., Mattei,J.-F., Rey,J., Rey,F. and Munnich,A. (1989) Genomics, **5**, 936–939.
14. Sommer,R. and Tautz,D. (1989) Nucleic Acids Res., **17**, 6749.
15. Bottema,C.D.K., Michels,V.V., Fisch,R.O. and Sommer,S.S. (1990) Amplifications, **4**, 27–29.
16. Nassal,M. and Rieger,A. (1990) Nucleic Acids Res., **18**, 3077–3078.
17. Petruska,J., Goodman,M.F., Boosalis,M.S., Sowers,L.C., Cheong,C. and Tinoco,I. (1988) Proc. Natl. Acad. Sci. USA, **85**, 6252–6256.
18. Verlan-de-Vries,M., Bogaard,M.F., van den Elst,H., van Boom,J.H., van der Eb,A.J. and Bos,J.L. (1986) Gene, **50**, 313–320.
19. Ehlen,T. and Dubeau,L. (1989) Biochem. Biophys. Res. Commun., **160**, 441–447.
20. Old,J.M., Varawalla,N.Y. and Weatherall,D.J. (1990) Lancet, **336**, 834–837.
21. Lichter-Konecki,U., Konecki,D.S., DiLella,A.G., Brayton,K., Marvit,J., Hahn,T.M., Trefz,F.K. and Woo,S.L.C. (1988) Biochemistry, **27**, 2881–2885.