Characterization of frequent polymorphisms in intron 2 of CYP21: application to analysis of segregation of CYP21 alleles

ANTHONY A. KILLEEN,1* RENÉE R. JIDDOU,1 and KUMUD S. SANE2

The gene encoding adrenal steroid 21-hydroxylase, CYP21, is located in the MHC class III region. Most cases of congenital adrenal hyperplasia (CAH) are caused by mutations in this gene, and most mutations appear to arise from gene conversion-like events involving the transfer of deleterious sequences from the pseudogene, CYP21P, which is located within 30 kb of CYP21. Approximately 20–30% of mutations are caused by deletions of CYP21. The second intron of CYP21 is polymorphic, and several base substitutions that include nt395, nt453, and nt601 have been reported; however, the frequencies of these polymorphisms are unknown. Using a combination of cleavage fragment length polymorphism analysis and direct sequencing, we examined the sequence of intron 2 in seven wild-type CYP21 genes and determined the frequency of polymorphisms at nt395, nt453, and nt601 in 48 chromosomes from the parental generation of Centre d’Etude du Polymorphisme Humain families. The observed frequencies of bases at these positions were as follows: 395C, 0.17; 395T, 0.83; 453C, 0.71; 453T, 0.29; 601A, 0.1; and 601C, 0.9. Using a PCR/restriction digestion approach to examine these intragenic markers, we could follow the segregation of alleles in informative families with 21-hydroxylase deficiency and identify deletions of CYP21. We emphasize that this method should be used in conjunction with other molecular genetic techniques for diagnosis of CAH. In addition to their potential use in families with CAH, these markers may be of use in genetic studies of the MHC in humans.
**Materials and Methods**

**SAMPLE PREPARATION AND PCR**

Blood samples were obtained from patients with congenital adrenal hyperplasia attributable to 21-hydroxylase deficiency and their relatives. Informed consent was obtained from all patients (or parents in the case of minors), and the study was approved by institutional review boards. The extraction of DNA from these samples has been described previously (11). Subjects with heterozygous deletions of CYP21 and the neighboring C4B gene were identified by Southern analysis using TaqI- and KpnI-restricted DNA probed with a 21-hydroxylase probe (12).

A nested PCR strategy was used to amplify portions of CYP21. The use of a nested technique was found to yield PCR products that were purer and quantitatively more consistent than those attainable with single PCR reactions. In addition, several of the DNA samples had been stored for up to 10 years and showed some degradation. These could be successfully amplified using a nested approach.

The first-round reaction used primers 1F and 832R, which amplify sequences from both CYP21 and CYP21P; the second-round reaction used primers 335F and b73R. Primer b73R recognizes the 8 bp in exon III of CYP21 that are deleted in CYP21P, and therefore can be used to selectively amplify the former gene. This primer has been described previously by Day et al. (13). For experiments involving cleavage fragment length polymorphism (CFLP) analysis, a 5’-biotinylated version of primer 335F was used for subsequent chemiluminescent detection of CFLP bands. The sequences and locations of hybridization of these primers are shown in Table 1. First-round PCR reactions were performed for 30 cycles with the following reagent concentrations: 20 mmol/L Tris-HCl, pH 8.4, 50 mmol/L KCl, 1.5 mmol/L MgCl2, 200 μmol/L each dNTP, 200 nmol/L each primer, 50–100 ng genomic DNA, and 2.5 U of Taq polymerase (Life Technologies) in a reaction volume of 50 or 100 μL. For the second-round reactions, PCR products from the first round were diluted 1000-fold with DNase-free water (Sigma Chemical Co.), and 1 μL of this dilution was used as template for the second round (volume, 100 μL), using the same reagent concentrations for 25 cycles. Annealing conditions were 60 °C for round 1 and 69 °C for round 2 reactions, both for 30 s. Extension conditions were 72 °C for 2 min (round 1) or 1 min (round 2). The latter reactions used a manual “hot-start” technique.

**CFLP**

For experiments involving CFLP analysis, exonuclease I digestion was performed to hydrolyze unincorporated primers and other single-stranded DNA present in the second-round PCR products. The magnesium concentration was increased to 5 mmol/L, and then 10 U of exonuclease I (Amersham Life Sciences) was added per microliter of PCR solution. The solution was incubated for 30 min at 37 °C; the exonuclease I was then inactivated by heating to 70 °C for 30 min. Finally, the PCR products were purified on High Pure PCR purification columns from Boehringer Mannheim, using the manufacturer’s protocol except that the final elution of DNA was with DNase-free water rather than the supplied buffer.

Cleavage I reactions were performed using kits from Third Wave Technologies. Reaction conditions were as follows: 10 mmol/L 2-(N-morpholino)propanesulfonic acid, pH 7.5, 0.5 g/L Tween 20, 0.5 g/L Nonidet P-40, 0.4 mmol/L MnCl2, and 50–100 fmol of DNA in a 10-μL reaction volume. The DNA was denatured by heating to 95 °C for 20 s, and then cooled to 55 °C. After 15 s, 2.5 μL (62.5 U) of cleavage I was added, and the reaction was allowed to proceed for 1 min, at which time it was stopped by the addition of 8 μL of stop solution consisting of 950 mL/L deionized formamide, 10 mmol/L EDTA, pH 8.0, 0.5 g/L xylene cyanol, and 0.5 g/L bromphenol blue.

Electrophoresis of CFLP fragments was performed in 8% denaturing polyacrylamide gels, 15 cm × 17 cm × 0.5 mm, containing 0.5× Tris-borate-EDTA buffer and 8 mol/L urea. Gels were prerun at 20 W for 30–60 min. CFLP products were denatured at 90 °C for 90 s, and then 9 μL was loaded in each lane. GelMarker I biotinylated molecular weight markers (Research Genetics) in the range of 50–700 bp were diluted 10-fold in loading dye, and 5 μL was loaded after heat denaturation. Gels were electrophoresed for 55 min at 20 W so that the xylene cyanol tracking dye had migrated ~90% of the length of the gel. Transfer of DNA to Nytran Plus nylon membrane (Schleicher & Schuell) was performed by capillary transfer overnight.

After DNA transfer, the membrane was blocked twice for 15 min using 100 mL of blocking solution (Amersham) on a rocking platform. Streptavidin-alkaline phosphatase conjugate [4 μL of 1000 U/mL stock solution (Amersham)] was diluted in 100 mL of blocking solution and applied to the membrane for 15 min with rocking. The membrane was then washed three times with SAAP buffer (100 mmol/L NaCl, 50 mmol/L Tris base, pH 10.0) containing 1 g/L sodium dodecyl sulfate for 5 min, and

---

**Table 1. Sequences and locations of primers used for PCR.**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’ → 3’)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1F</td>
<td>ATG CTG CTC CTG GCC CTG C</td>
<td>Exon I</td>
</tr>
<tr>
<td>832R</td>
<td>CTC ACA GAA CTC CTG GGT CAG CTG CTC</td>
<td>Exon III</td>
</tr>
<tr>
<td>(Bio-)335F</td>
<td>(Biotin-)GAA GCC ATG CTC AAA AAG TGG GGG CTT TCC AGA GCA GGA AGT AGT</td>
<td>Exon II</td>
</tr>
<tr>
<td>b73R</td>
<td>GGG CTT TCC AGA GCA GGA AGT AGT CT</td>
<td>Exon III</td>
</tr>
</tbody>
</table>
then three times with SAAP buffer containing 1 mmol/L MgCl₂ for 5 min. After the final wash, the membrane was placed in a clear polythene bag (Oncor). SAAP buffer (3 mL) containing 30 μL of 100× CDP-Star (Boehringer Mannheim) was added and allowed to coat the membrane for 3 min. The excess buffer was discarded, and the bag was sealed. Chemiluminescent bands were then visualized by placing the bag directly against Kodak XOMAT-AR film in the dark for 30 min, after which the film was developed.

**DNA SEQUENCING**

DNA sequencing of the purified PCR products prepared for the CFLP reactions was performed with an ABI 373 sequencer using standard methodologies.

**PCR/RESTRICTION ENZYME DIGESTION**

Second-round PCR products (10 μL) were digested with either 2.5 U of HaeIII (Life Technologies) or 2 U of BbvI (New England Biolabs) after addition of 1 μL of the manufacturer’s 10× restriction buffer. After incubation for 2 h at 37 °C, restriction products were electrophoresed in 2% agarose gels containing 1 mg/L ethidium bromide and photographed on an ultraviolet transilluminator.

**Results**

**IDENTIFICATION OF FREQUENT POLYMORPHISMS**

DNA samples from seven clinically healthy subjects, previously determined to be heterozygotes for deletions of CYP21 as determined by Southern analysis of TaqI- and KpnI-restricted DNA, were initially tested for polymorphisms, using CFLP. These subjects have one wild-type CYP21 gene, which could be examined in isolation for polymorphisms. This strategy facilitates recognition of polymorphisms that are present on the same allele. The CFLP patterns obtained from DNA of these subjects showed two polymorphisms (Fig. 2). The first was the presence or absence of a doublet of CFLP bands located ~60 bp from the 5’ end of the sense strand. A second CFLP polymorphism, characterized by the absence of a CFLP band ~190 bp from the 5’ end of the sense strand, was seen in one subject. Sequence analysis demonstrated that the first CFLP polymorphism correlated with a pair of nucleotides that are in linkage disequilibrium in the samples studied: C/T at nt395 and A/C at nt419, the two bands being present when the nucleotides at these positions are, respectively, C and A, and absent when they are T and C. The second CFLP polymorphism correlated with two nucleotide substitutions at nt601 and nt605.

Sequence analysis showed several additional polymorphisms. Of these, one was seen in more than two subjects: C/T at nt453 (Fig. 3). A deletion of G at nt564 was seen in two subjects. Of all the different polymorphisms, three occur at restriction enzyme recognition sites: nt395 and nt601 are in HaeIII restriction sites; and nt453 is in a BbvI recognition site. There are also additional restriction sites for both of these enzymes within the amplified region: one HaeIII site is near nt614; and one BbvI recognition site is near nt623. Restriction with BbvI gives rise to a constant fragment of 125 bp and polymorphic fragments of 278 bp when nt453 is T or 169 and 109 bp when nt453 is C. It should be noted that the restriction cutting site for BbvI is offset from the recognition site. The BbvI recognition site partially overlaps the previously described PvuII recognition site, which is rarely polymorphic (10). Restriction with HaeIII produces a constant fragment of 118 bp and polymorphic fragments of either 265 bp when nt395 is T or 206 and 59 bp when nt395 is C. If the polymorphic HaeIII site at nt601 is also mutated, the latter fragments are 15 bp longer.

**FREQUENCIES OF POLYMORPHISMS**

The frequencies of the HaeIII and BbvI polymorphisms, which fall on restriction sites, were determined in DNA samples from 24 Centre d’Étude du Polymorphisme Humain (CEPH) subjects in the parental generation (48 chromosomes). The frequency of each polymorphism was as follows: 395C, 8 of 48; 395T, 40 of 48; 453C, 34 of 48; and 453T, 14 of 48. The frequencies of the genotypes at these positions are shown in Table 2. Values of χ² for the observed and expected distributions of genotypes for the nt395 and nt453 polymorphisms were 3.83 and 0.23, respectively, (P >0.05 with 2 degrees of freedom for each polymorphism). The distributions are consistent with Hardy-Weinberg equilibrium. The nt601A polymorphism was seen in only 5 of 48 chromosomes.
Inheritance of the nt395 and nt453 polymorphic markers was studied in families with known deletions of CYP21. An illustrative analysis of these markers in an informative family with deletions of CYP21 is shown in Fig. 4. Each parent is heterozygous for the CYP21 deletion associated with the HLA-A3, B47, DR7 haplotype. The affected child, who is known to be homozygous for the deletion haplotypes, has not inherited markers of a CYP21 gene from either parent. The presence of available template DNA for the second-round PCR reaction was confirmed by gel analysis. The unaffected child has inherited polymorphic markers from each parent, demonstrating the presence of two distinct CYP21 alleles.

Inheritance of the nt395 and nt453 polymorphic markers was studied in families with known deletions of CYP21. An illustrative analysis of these markers in an informative family with deletions of CYP21 is shown in Fig. 4. Each parent is heterozygous for the CYP21 deletion associated with the HLA-A3, B47, DR7 haplotype. The affected child, who is known to be homozygous for the deletion haplotypes, has not inherited markers of a CYP21 gene from either parent. The presence of available template DNA for the second-round PCR reaction was confirmed by gel analysis. The unaffected child has inherited polymorphic markers from each parent, demonstrating the presence of two distinct CYP21 alleles.

**Discussion**

Using a combination of CFLP, direct sequencing, and CEPH family screening with a PCR/restriction digestion approach, we have shown that nt395, nt419, and nt453 in the second intron of CYP21 are frequently polymorphic. A polymorphism at nt601 is less frequent, being found in only 5 of 48 (10%) chromosomes in the CEPH families studied. Linkage disequilibrium between polymorphisms at nt395 and nt419 was observed in seven wild-type alleles: 395T was found in conjunction with 419C; and 395C was found in conjunction with 419A. However, the
CYP21 alleles in families affected by 21-hydroxylase deficiency. In general, direct mutation detection is preferred for molecular diagnostic studies because methods based on the linkage of polymorphic markers are subject to errors arising from recombination events, new mutations, non-paternity, and mistaken diagnosis in a proband, and cannot be applied if a pedigree is uninformative for the markers tested. However, in the case of CYP21 mutation analysis, the presence of the highly homologous CYP21P complicates the routine application of direct mutation testing. To selectively amplify the CYP21 gene for analysis by PCR, gene-specific primers must be used. Because of the high frequency of gene conversions from CYP21P to CYP21, critical sequences for CYP21 gene-specific amplification may be altered, preventing amplification of the expected gene. Furthermore, frequent “back conversions” of CYP21 sequences to CYP21P have been reported (13). Selection of a primer sequence that coincides with such a mutation might lead to undesired amplification of portions of CYP21P and thus, theoretically at least, to the possibility of mistaken identification of a mutation. In addition, it may not always be possible to identify a mutation in affected individuals by direct mutation testing, and several investigators have used other polymorphic markers within the MHC, such as HLA typing or MHC microsatellites, when evaluating relatives of a known patient with congenital adrenal hyperplasia (16). Analysis of the segregation pattern of the polymorphisms described in the study may provide additional useful genetic information when evaluating relatives of an affected patient.

These polymorphisms may also be useful in identifying deletions of CYP21P, which account for ~20–30% of all mutations (6). The currently preferred method to detect deletions of CYP21 is to perform Southern analysis on TaqI and BglII-restricted DNA, using a 21-hydroxylase probe. Relative gene dosages of CYP21P and CYP21 can then be assessed by densitometry. If a family is informative for the polymorphic markers described in this study, a gene deletion leads to the absence of parental markers in an affected offspring. However, one must bear in mind that detection of these markers depends on the successful amplification of the CYP21P alleles. The downstream primer b73R is specific for the 8 bp in CYP21P that are deleted in CYP21P. Gene conversion events that transfer this deletion to CYP21 produce a mutant allele that would not be expected to amplify when the PCR approach described here is used. This pattern might be confused with a CYP21 deletion. An alternative CYP21-specific primer could be used if the family was known or suspected to have a gene deletion, for example, from HLA typing or Southern analysis. Other investigators have reported CYP21-specific primers (13, 17). Furthermore, the occasional dropout of alleles during PCR amplification has been observed in intron 2 of CYP21 (18). For these reasons, the markers described here must be used cautiously, and in conjunction with other molecular genetic studies, for diagnostic purposes.

![Fig. 4. Analysis of segregation of polymorphic BbvI and HaeIII restriction sites in a family with congenital adrenal hyperplasia.](Image)

The father and mother are heterozygotes for CYP21 deletions associated with the MHC haplotype that includes HLA-A3, B47, DR7. The sequences of their remaining alleles are shown in Fig. 3 (subjects O6 and O7, respectively). The affected child has not inherited a CYP21 allele from either parent. In contrast, the unaffected child has inherited markers from both parents.

markers at nt395 and nt453 are not in complete linkage disequilibrium, as shown by the analysis of wild-type alleles (Fig. 3). It is therefore of value to test both these markers when screening for polymorphisms. The C/T transition at nt395 occurs at a CG dinucleotide, a frequent context for this kind of mutation (14).

The linked polymorphisms at nt395/nt419 and those at nt601/nt605 were initially observed by CFLP analysis. The polymorphism at nt453 did not alter the CFLP patterns under the experimental conditions used, nor did the single nucleotide polymorphism (A/G) at nt683, which produces a lysine/arginine (AAG/AGG) codon substitution in exon III. The detection of polymorphisms by CFLP depends on the recognition and cleavage of sequence-specific secondary structure in the target DNA (15). It may be possible to detect these polymorphisms by CFLP analysis of the antisense strand or with the use of other cleavage enzymes. However, this issue was not investigated further because the frequent nt395 and nt453 polymorphisms can be detected easily with commercially available restriction enzymes, using a PCR/restriction enzyme digestion technique. For both of these enzymes, there are also constant restriction sites in the amplified region of CYP21, which can serve as internal controls for restriction endonuclease cutting.

These polymorphisms are sufficiently frequent to be of use in following the segregation of CYP21 alleles in genetic studies. They can also be used to follow the segregation of mutant CYP21 alleles in families affected
The data presented in this study confirm the high frequency of polymorphism found in intron 2 of CYP21. These markers may be of use in studying the mechanisms of gene conversion and in determining the extent of gene conversion events, which appear to give rise to the majority of mutations in 21-hydroxylase deficiency alleles. In addition, because of their location in the class III region of the HLA complex, these markers may be of use in studies of the genetics of the human MHC.

References