ORIGINAL ARTICLE

Hsien-Hsiung Lee · Dau-Ming Niu · Ruey-Wen Lin Peter Chan · Ching-Yu Lin

Structural analysis of the chimeric *CYP21P/CYP21* gene in steroid 21-hydroxylase deficiency

Received: March 13, 2002 / Accepted: June 17, 2002

Abstract Congenital adrenal hyperplasia (CAH) is a common autosomal recessive disorder mainly caused by defects in the steroid 21-hydroxylase (CYP21) gene. More than 90% of CAH cases are caused by mutations of the CYP21 gene. Approximately 75% of the defective CYP21 genes are generated through intergenic recombination, termed "apparent gene conversion," from the neighboring CYP21P pseudogene. A chimeric CYP21P/CYP21 gene with its 5' end corresponding to CYP21P and 3' end corresponding to CYP21 has been identified. This type of gene is nonfunctional because it produces a truncated protein. We found two distinct chimeric genes in CAH patients. Both genes had a sequence with -300 nucleotides of the 5' head as the CYP21P gene. The coding region consisted of a fusion molecule with the CYP21P gene in two different regions. One of the junctions was located in the chi-like sequence of GCTGGGC in the third intron and the other was in the minisatellite consensus TGGCAGGAGG of exon 5 of the CYP21P gene. In addition, analysis of restriction fragment length polymorphism for these two 3.3-kb chimeric molecules showed that these sequences arose as a consequence of unequal crossover between the CYP21P and CYP21 genes. It is plausible that both consensus sequences are responsible for the gene conversion of these two chimeric genes.

Key words Chimeric gene \cdot Unequal crossover \cdot Chi-like and minisatellite sequences \cdot Congenital adrenal hyperplasia \cdot CYP21 deficiency

King Car Food Industrial Co., Ltd., Yuan-Shan Research Institute, #326, Sec. 2, Yuan-Shan Rd., Yuan San, I-Lan 264, Taiwan, Republic of China

Tel. +886-3-9229000; Fax +886-3-9228030

e-mail: hhlee@ms2.kingcar.com.tw

D.-M. Niu

Introduction

The gene coding for P450c21 is designated CYP21. More than 90% of congenital adrenal hyperplasia (CAH) cases are caused by mutation of the CYP21 gene (White et al. 1987; Miller and Morel 1989). There exists a duplicated copy close by that is designated CYP21P, which is a pseudogene and shares 98% nucleotide sequence homology with CYP21 in the exon sequences (White et al. 1986; Higashi et al. 1986; Rodrigues et al. 1987). These two genes are about 30kb apart and are located within the human leukocyte antigen (HLA) class III human histocompatibility complex locus on chromosome 6p21.3, adjacent to the C4A and C4B genes encoding the fourth components of serum complement (White et al. 1984). Because of a high recombination rate in the human major histocompatibility complex and the CYP21 gene being located within the HLA complex, the frequency of recombination events between CYP21P and CYP21 genes is relatively high in the human population. Occasionally, 1 of the 15 mutations found in the CYP21P gene is inserted into the CYP21 gene by "gene conversion mutations," the mechanism of which is not well understood (Donohoue et al. 1986). On the other hand, gross gene deletions (Caroll et al. 1985; Rodrigues et al. 1987; White et al. 1988; Speiser et al. 1991) encompassing the C4A and CYP21 genes as a result of a 30-kb deletion by unequal crossover in meiosis have been reported to occur in 20%–30% of alleles in most populations (White et al. 1988). However, such deletions probably represent gene conversion, unequal crossover, or polymorphisms rather than simple gene deletion (Matteson et al. 1987; Miller 1988).

The chimeric *CYP21P/CYP21* gene with its 5' and 3' ends corresponding to *CYP21P* and *CYP21* (Tuie-Luna and White 1995), respectively, is nonfunctional, because there is a deleterious mutation in G110delGA-Y112del, which results in a frameshift and produces a truncated protein. This leads to complete abolishment of enzyme activity of steroid 21-hydroxylase. In the present study, we report two chimeric *CYP21P/CYP21* genes from CAH patients in our population. Sequence analysis showed that both had a se-

H.-H. Lee $(\boxtimes) \cdot R.-W.$ Lin $\cdot P.$ Chan $\cdot C.-Y.$ Lin

Department of Pediatrics, Veterans General Hospital-Taipei, Taiwan, Republic of China

quence with over -300 nucleotides of the 5' head of the *CYP21P* gene. The coding region consisted of a fusion molecule with the *CYP21P* gene in two different regions. One of the junctions was located in the chi-like sequence GCTGGGC in the third intron and the other was in the minisatellite consensus TGGCAGGAGG of exon 5 of the *CYP21P* gene. Both consensus sequences may be responsible for gene conversion, and these two chimeric *CYP21P*/*CYP21* genes result from a 30-kb deletion in the region of HLA class III (L' Allemand et al. 2000).

Materials and methods

Subjects

As previously reported (Lee et al. 2000a), we found two individuals who carried a chimeric *CYP21P/CYP21* gene in one of the chromosomes with 21-hydroxylase deficiency. The Institute Review Board of Veterans General Hospital-Taipei approved the protocol, and the study strictly followed their guidelines (Lee et al. 2000b).

Amplification and cloning of the chimeric CYP21P/ CYP21 gene

To amplify the chimeric *CYP21P/CYP21* genes, we mixed BF1, AF1, and 21BR primers (Lee et al. 2000a) (Table 1) in a 50-µl reaction and polymerase chain reaction (PCR) was performed in a mixture of *TAQ/POW* polymerases (Roche Diagnostics, Mannheum, Germany) under conditions described previously (Lee et al. 2000a). Fragments of 3.5 kb (Lee et al. 2000a; Lee 2001) covering the sequences of the *CYP21* genes from -313 to 3170 (Higashi et al. 1986) were generated. The PCR products were further subcloned to the pGEM-T vector system (Promega, Madison, WI, USA).

Analysis of the chimeric *CYP21P/CYP21* gene by the amplification-created restriction site method (ACRS)

After transformation, the colonies were screened by PCR amplification with the ACRS method (Lee et al. 1996). ACRS primer pairs used for mutation detection in P30L (CIN/C2), nucleotide (nt) 655 (C3B/C4A), and G110delGA-Y112del (C3B/C4A) are listed in Table 1. For further analysis of mutations in I172N, I236N, V281L, Q318X, and R356Q, the ACRS primers and reaction conditions have been described previously (Lee et al. 1996). Mutational detection in the promoter region and intron sequence was carried out by DNA sequencing.

Analysis of the chimeric *CYP21P/CYP21* gene by restriction fragment length polymorphism (RFLP)

DNA of pGEM-T subclone containing mutations at P30L, nt 655, and G110delGA-Y112del was purified and reamplified with primers AF1/21BR. The normal *CYP21* and *CYP21P* genes were also generated from a normal individual genome with primers BF1/21BR and AF1/21AF, respectively (Table 1). The reaction conditions have been described previously (Lee et al. 1996). Each 3.5-kb PCR product was subject to digestion with *AfI*II, *AseI*, and *AlwI* in a 20-µl reaction volume at 37°C overnight. After complete digestion, the reaction mixture was analyzed on agarose gel.

DNA sequencing

The PCR product amplified from identified clones was directly used as a template for DNA sequencing. The reaction mixture of the ABI PRISM Dye Termination Cycle Sequencing Kit (Perkin-Elmer, Foster City, CA, USA) was applied to the DNA cycling reaction. The protocol used was as described previously (Lee et al. 1998).

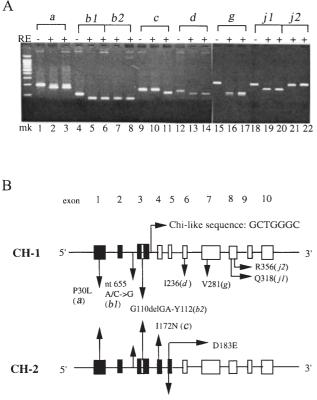
Table 1. Primers for the analysis of the CYP21P/CYP21 chimeric molecule

Designation	Primer $(5' \rightarrow 3')$	Location (nt) ^a	Specificity					
Amplification of the CYP21P/CYP21 chimeric molecule								
BF1	CCCAGGTGGGGGGGGGACACTA	-314 to -294	CYP21					
21BR	AATTAAGCCTCAATCCTCTGCAGCG	3177 to 3153	CYP21					
AF1	CCCAGGTCGGGGCGGACACCC	-315 to -295	CYP21P					
21AR	GATTAAGCCTCAATCCTCTGCGGCA	3185 to 3161	CYP21P					
Amplification for ACRS analysis of specific regions								
CIN	CTACACAGCAGGAGGGGGATGGC	-53 to -63	P30					
C2	AGCAAGTGCAAGAAGCCCGGGGCAAGctG ^b	122 to 94	P30					
C3B	TTCATCAGTTCCCACCTCCAGCCCCgA	631 to 658	nt 655					
C4A	CTTCTTGTGGGCTTTCCAGAGCAGgtA	759 to 740	G110delGA-Y112del					
C5	GAGGAATTCTCTCTCCTCACCTGCĂGCATtA	970 to 1003	I172					
C6	TTGTCGTCCTGCCAGAAAAGGA	1131 to 1110	I172					

ACRS, Amplification-created restriction site method; nt, nucleotide; del, deletion

^aNucleotide sequence based on Higashi et al. (1986)

^bLowercase letters indicate modified nucleotides. C2, C3B, C4A, and C5 primers create *PstI* (CTGCAG), *SacI* (GAGCTC), *RsaI* (GTAC), and *MseI* (TTAA) sites, respectively, whereas there is a base change in the P30, nt 655 A/C→G, G110delGA-Y112del, and I172 (Lee et al. 1996)



Minisatellite consensus: TGGCAGGAGG

Fig. 1A,B. Restriction analysis of the amplification-created restriction site method (ACRS) amplification product and map for detection of mutational loci of the chimeric genes. The primer pairs for the detection of P30L, nucleotide (nt) 655, G110delGA-Y112del, and I172N are shown in Table 1. Other mutations detected by ACRS primers have been described previously (Lee et al. 1996). DNA for ACRS primer amplification was used from pGEM-T identified clones. A Analysis on 2.5% Metaphor (FMC Bioproducts, Rockland, ME, USA). Detection of ACRS amplification products from each locus are designated as a for P30L, b1 for nt 655, b2 for G110delGA-Y112del, c for I172N, d for I236N, g for V281L, j1 for Q318X, and j2 for R356Q. Lanes 2, 5, 7, 10, 13, 16, 19, and 21 were for clone CH-1 analysis. Lanes 3, 6, 8, 11, 14, 17, 20, and 22 were for clone CH-2 analysis. Each amplification product was either untreated (minus signs) or treated (plus signs) with an appropriate restriction enzyme (RE) (Lee et al. 1996). The marker used was a 100-bp ladder in lane mk. B Schematic representation of the detection of ACRS for CH-1 and CH-2 clones. Exons for CYP21P and CYP21 genes are shown as solid blocks and open boxes, respectively. The arrows point to the position of the locus corresponding to those designated in A

Results

Isolation and identification of the chimeric *CYP21P/CYP21* gene

Two CAH families carrying the chimeric *CYP21P/CYP21* gene were analyzed (Lee et al. 2000a). A total of 3.5kb of *CYP21* DNA generated by allele-specific primers (BF1/AF1/21BR) was purified with agarose gel and subcloned to the pGEM-T vector system (Promega). Two clones obtained separately from two families, CH-1 (Fig. 1A, lanes 2, 5, and 7) and CH-2 (Fig. 1A, lanes 3, 6, and 8), identified by

the ACRS primer (Table 1) had mutations in P30L (Fig. 1A,a), nt 655A/C→G (Fig. 1A,b1), and G110delGA-Y112del (Fig. 1A,b2). From further analysis by the ACRS method (Fig. 1A), CH-1 clone had no other mutation found in other loci such as I172N (Fig. 1A,c, lane 10), I236N (Fig. 1A,d, lane 13), V281L (Fig. 1A,g, lane 16), Q318X (Fig. 1A,j1, lane 19), and R356Q (Fig. 1A,j2, lane 21). However, CH-2 carried a mutation in codon 172 (Fig. 1A,c, lane 11) and lacked mutations at I236N (Fig. 1A,d, lane 14), V281L (Fig. 1A,g, lane 17), Q318X (Fig. 1A,j1, lane 20), and R356Q (Fig. 1A,j2, lane 22). For further DNA sequencing analysis as shown in Table 2, CH-1 had no mutation in D183E, which exists in the CYP21P gene. CH-2 carried a mutation in this locus as in the CYP21P gene. In addition, CH-1 was identical to the CYP21P gene from exon 1 to exon 3 (Fig. 1B) and CH-2 was identical to the CYP21P gene from exon 1 to exon 5 (Fig. 1B). Interestingly, in nt 395, CH-1 has a C base instead of a T, whereas CH-2 contains a T base (Table 2). These two clones had normal alleles in codon 306 (data not shown).

RFLP analysis for the chimeric CYP21P/CYP21 gene

The 3.5-kb fragments amplified with primers AF1/21BR from CH-1 and CH-2 were digested with AffII, AlwI, and AseI restriction endonucleases. The CYP21P gene from a normal individual was amplified with primers AF1/21AR and the CYP21 gene was amplified with primers BF1/21BR (Fig. 2A, lane C). From restriction analysis (Fig. 2A) and restriction mapping (Fig. 2B), the CYP21P gene (Fig. 2A, lane 1) can be digested into four fragments of 2599, 688, 115, and 85 bp caused by cleavage with AfIII at nt 3087, AlwI at nt 497, and AseI at nt -202 (Higashi et al. 1986) (Fig. 2B). The normal CYP21 gene (Fig. 2A, lane 2) shows 1705-, 898-, 796-, and 85-bp fragments, as a result of cleavage by AffII at nt 3085 and AlwI at nt 488 and 1372 (Higashi et al. 1986) (Fig. 2B). There is no AseI cleavage site in the CYP21 gene (Higashi et al. 1986). Both CH-1 and CH-2 genes were digested into 1705-, 899-, 696-, 115-, and 85-bp fragments (Fig. 2A, lanes 3 and 4), as a result of the cleavage of AffII at nt 3087, AseI at nt -202, and AlwI at nt 494 and 1372 (Fig. 2B). Both 85-bp and 115-bp fragments were run out of the gel during electrophoresis. Both CH-1 and CH-2 clones were two distinct chimeric genes caused by an unequal crossover reaction.

Discussion

About 20% of alleles in 21-hydroxylase deficiency have a 30-kb deletion including the 3' end of *CYP21P*, all of *C4B*, and the 5' end of the *CYP21* gene. This is presumably due to unequal crossover during meiosis (L'Allemand et al. 2000). The precise breakpoint in the 5' head and the 3' end of the *CYP21P–CYP21* transition area was never determined because of the long sequence identity between *CYP21P* and *CYP21*. However, some studies have sug-

Sequence* (nt/an) of $CYP2I$ $CYP2IP$ $CH-1$ $CH-2$ nt -306 G C C C C -2842 A G G G C -282 A G C C C -198 C T C C C -188/189 - +T +T +T +T -126 C T T T T -113 G A A A A -110 T C C C C -103 A G G G G -100 P L L L L P30 P C C T T A4 C C C C T 431/432 - +TGTT +TGTT +TGTT +TGTT 431/422 T C C				Types of chimeric gene	
-284/-293 T/A C/C C/C C/C C/C -289 A G G G G -209 T C C C C -198 C T C C C -188/189 +T T T T -110 T C C C C -110 T C C C C -103 A G A A A -100 P L L L L 239 L(TTG) L(CTG) L(CTG) L(CTG) 1395 T C C T 413432 +TGTT +TGTT +TGTT 420 A G G G G 431/432 +TGTT +TGTT +TGTT +TGTT 470 A G G G G <th>Sequence^a (nt/aa) of</th> <th>CYP21</th> <th>CYP21P</th> <th>CH-1</th> <th>CH-2</th>	Sequence ^a (nt/aa) of	CYP21	CYP21P	CH-1	CH-2
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
-126 C T T T T -113 G A A A -110 T C C C -103 A G G G -104 C T T T P10 P L L L P30 P L L L 139 L(TTG) L(CTG) L(CTG) P(CCC) P(CCA) P(CCC) P(CCC) P(CCC) P(CCC) x395 T C C T TGTT +TGTT 412 T G G G G G G 413/452 +TGTT +TGTT +TGTT +TGTT 41GTT 412 T G G G G G G 513 T G G G G G G G 524/511 +TCC T T T T G A A A <		С			
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		—			
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	-126	С	Т	Т	Т
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	-113	G		А	А
-4 C T T T T P10 P L L L P30 P L L L L39 L(TTG) L(CTG) L(CTG) P(CC) P45 P(CCA) P(CCC) P(CCC) P(CCC) n395 T C C T 431/432 +TGTT +TGTT +TGTT 442 T G G G G 470 A G G G G G 470 A G G G G G 470 A G G G G G 513 T G G G G G 524 T C C C C C 540/541 - +TCC +TCC +TCC S T 540/551 GA A <td< td=""><td>-110</td><td>Т</td><td>С</td><td>С</td><td>С</td></td<>	-110	Т	С	С	С
P10 P L L L L P30 P L L L L P30 P L CTG L(CTG) L(CTG) L39 L(TTG) L(CTG) P(CCC) P(CCC) P45 P(CCA) P(CCC) P(CCC) P(CCC) r1395 T C C T 431/432 +TGTT +TGTT +TGTT 442 T G G G 470 A G G G 470 A G G G G 471 T G G G G 502 A G G G G 513 T G G G G 524 T C C C C 540/541 - +TCC +TCC +TCC +TCC 577 G A A A A 589 G T T	-103	А		G	G
P30 P L L L L L39 L(TTG) L(CTG) L(CTG) L(CTG) P45 P(CCA) P(CCC) P(CCC) P(CCC) nt 395 T C C T 431/432 +TGTT +TGTT +TGTT +TGTT 442 T G G G G 470 A G G G G 470 A G G G G 478 A G G G G 502 A G G G G 513 T G C C C 524 T C C C T 530 A T T T T 540/541 +TCC +TCC +TCC +TCC 572-574 GGT T T T T 580/581 G/A A/G A/G A A	-4	С	Т	Т	Т
L39 L(TTG) L(CTG) L(CTG) L(CTG) L(CTG) P45 P(CCA) P(CCC) P(CCC) P(CCC) P(CCC) nt 395 T C C T T 431/432 +TGTT +TGTT +TGTT +TGTT 442 T G G G G 470 A G G G G 502 A G G G G G 513 T C C C C C 540/541 - +TCC +TCC +TCC +TCC 577 G A A/G A/G A/G 580/581 G/A <	P10	Р	L	L	L
P45 P(CCA) P(CCC) P(CCC) P(CCC) P(CCC) nt 395 T C C T 431/432 +TGTT +TGTT +TGTT 442 T G G G G 470 A G G G G 478 A G G G G 486 A G G G G 502 A G A G G 513 T G G G G 524 T C C C C 530 A T T T T 540/541 +TCC +TCC +TCC +TCC 546 C T T T T T 577 G A A/G A/G A/G A/G 580/581 G/A A/G	P30	Р	L	L	L
P45 P(CCA) P(CCC) P(CCC) P(CCC) P(CCC) nt 395 T C C T 431/432 +TGTT +TGTT +TGTT 442 T G G G G 470 A G G G G 478 A G G G G 486 A G G G G 502 A G A G G 513 T G G G G 524 T C C C C 530 A T T T T 540/541 +TCC +TCC +TCC +TCC 546 C T T T T T 577 G A A/G A/G A/G A/G 580/581 G/A A/G	L39	L(TTG)	L (CTG)	L(CTG)	L(CTG)
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	P45		P(CCC)	P(CCC)	P(CCC)
442 T G G G G 470 A G G G 478 A G G G 478 A G G G 486 A G G G 502 A G G G 513 T G G G 524 T C C C 530 A T T T 540/541 +TCC +TCC +TCC 540 C T T T T 540/541 +TCC TCA TCA TCA 572-574 GGT TCA TCA A/G A/G 589 G T T T T 589 G T T T T 594 G A A A A 610 C A G G G 624 G <t< td=""><td>nt 395</td><td></td><td>C</td><td>C</td><td></td></t<>	nt 395		C	C	
470 A G G G G 478 A G G G 486 A G G G 502 A G A G 513 T G G G 513 T C C C 530 A T T T 540/541 +TCC +TCC +TCC 546 C T T T T 540/541 +TCC +TCC +TCC 54 577 G A A A A 577 G A A A A 580/581 G/A A/G A/G A/G A 59 G T T T T 5 594 G A A A A 6 610 C A G G G G G G G G G G	431/432	_	+TGTT	+TGTT	+TGTT
478 A G G G G 486 A G G G G 502 A G A G G 513 T G G G G 524 T C C C T 530 A T T T T 540/541 +TCC +TCC +TCC +TCC 546 C T T T T 572-574 GGT TCA TCA A A 577 G A A A A 589 G T T T T 594 G A A A A 601 C A A A A 610 C A G G G G 629 A G G G/G G/G G/G G/G 610 C/A G/G G/G G/G <td>442</td> <td>Т</td> <td>G</td> <td>G</td> <td>G</td>	442	Т	G	G	G
486 A G G G G 502 A G A G 513 T G G G 513 T C C C 530 A T T T 540/541 +TCC +TCC +TCC 546 C T T T 572-574 GGT TCA TCA TCA 577 G A A A 577 G A A A 589 G T T T 594 G A A A 601 C A A A 610 C A A A 620 A G G G G 624 G T T T T 629(630 C/A G G G G 6100 G(GGA) G110delGA G110delGA G110delGA <	470	А	G	G	G
502 A G A G 513 T G G G 513 T C C C 524 T C C C 530 A T T T 540/541 +TCC +TCC +TCC 546 C T T T 572-574 GGT TCA TCA TCA 577 G A A A 580/581 G/A A/G A/G A/G 589 G T T T 594 G A A A 610 C A A A 624 G T T T 629/630 C/A G G G 610 C/A G G G 610 G(GGA) G10delGA G10delGA G10delGA 6110 G(GGA) G10delGA G10delGA G10delGA 6110 <td>478</td> <td>А</td> <td>G</td> <td>G</td> <td>G</td>	478	А	G	G	G
513 T G G G G 524 T C C C 530 A T T T 540/541 $-$ +TCC +TCC +TCC 540/541 $-$ +TCC +TCC +TCC 574 GGT TCA T T 572-574 GGT TCA A A 589 G A A A 589 G T T T 594 G A A A 601 C A A A 610 C A A A 624 G T T T 629 A/C G G G 620 A G G G G 655 A/C G G G G 6100 G(GGA) G110delGA G110delGA G110delGA 9111 D Del Del Del	486	А	G	G	G
513 T G G G G 524 T C C C 530 A T T T 540/541 $-$ +TCC +TCC +TCC 540/541 $-$ +TCC +TCC +TCC 574 GGT TCA T T 572-574 GGT TCA A A 589 G A A A 589 G T T T 594 G A A A 601 C A A A 610 C A A A 624 G T T T 629 A/C G G G 620 A G G G G 655 A/C G G G G 6100 G(GGA) G110delGA G110delGA G110delGA 9111 D Del Del Del	502	А	G	A	G
524 T C C C C 530 A T T T T 540/541 $ +$ TCC $+$ TCC $+$ TCC $+$ TCC 546 C T T T T 572-574 GGT TCA TCA TCA 577 G A A A 580/581 G/A A/G A/G A/G 589 G T T T 594 G A A A 601 C A A A 610 C A A A 620 A G G G 624 G T T T 629/630 C/A G/G G/G G/G 610 A G G G G 6110 G(GGA) G110delGA G110delGA G110delGA 6110 G(GGA) G110delGA G110delGA G10delGA 77 <td></td> <td></td> <td></td> <td></td> <td></td>					
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
540/541 +TCC +TCC +TCC +TCC 546 C T T T T $572-574$ GGT TCA TCA TCA 577 G A A A $580/581$ G/A A/G A/G A/G 589 G T T T 594 G A A A 601 C A G G 624 G T T T $629/630$ C/A G/G G/G G/G 610 A G G G G 6100 A G G G G 6100 G/G G/G G/G G/G G/G 6100 G G G G G G					
546 C T T T T $572-574$ GGT TCA TCA TCA 577 G A A A $580/581$ G/A A/G A/G A/G 589 G T T T 594 G A A A 601 C A A A 611 C A A A 6120 A G G G G 620 A G G G G 6103 C/A G/G G/G G G 6104 G G/G G/G G G 6104 G G/G G G G G 6104 G G/G G G G G G		_			
572-574 GGT TCA TCA TCA TCA 577 G A A A $580/581$ G/A A/G A/G A/G 589 G T T T 594 G A A A 601 C A A A 601 C A A A 610 C A A A 610 C A A A 610 C A A A 620 A G G G $629/630$ C/A G/G G/G G/G 620 A G G G G 620 A G G G G G 620 A G G G G G 61040 G/G G/G G G G G G 5108 S(TCC) S(TCC) S(TCT) S		С			
577 G A A A $580/581$ G/A A/G A/G A/G 589 G T T T 594 G A A A 601 C A A A 601 C A A A 601 C A A A 610 C A A A 610 C A A A 610 G T T T $629/630$ C/A G/G G/G G/G 620 A G G G G 620 A/C G G G G 610 C/A G G G G 610 A/C G G G G 610 G(GA) G110delGA G110delGA G110delGA G110delGA 6110 D Del Del Del Del S <					
580/581 G/A A/G A/G A/G 589 G T T T T 594 G A A A A 601 C A A A A 601 C A A A A 610 C A A A A 610 C A A A A 624 G T T T T $629/630$ C/A G/G G/G G/G G 620 A G G G G G 620 A G G G G G G 620 A G					
589 G T T T 594 G A A A 601 C A A A 601 C A A A 611 C A A A 611 G G G G 624 G T T T $629/630$ C/A G/G G/G G/G 620 A G G G G 620 A G G G G 655 A/C G G G G 6108 S(TCC) S(TCG) S(TCG) S(TCG) $G110$ G(GGA) G110delGA G110delGA G110delGA $D111$ D Del Del Del S(TCT) $S113$ S(TCC) S(TCT) S(TCT) S(TCT) 1860 C C C <					
594 G A A A 601 C A A A nt 620 A G G G 624 G T T T 629/630 C/A G/G G/G G/G 620 A G G G 620 A G G G 655 A/C G G G 5108 S(TCC) S(TCG) S(TCG) S(TCG) 6110 G(GGA) G110delGA G110delGA G110delGA D111 D Del Del Del S(TCT) Y112 Y Del Del Del S(TCT) S(TCC) S(TCT) S(TCT) S(TCT) S(TCT) 113 S(TCC) S(TCT) S(TCT) S(TCT) S(TCT) 1172 I N I N N 1172 I N I <td></td> <td></td> <td></td> <td></td> <td></td>					
601 C A A A nt 620 A G G G 624 G T T T $629/630$ C/A G/G G/G G/G 620 A G G G 620 A G G G 655 A/C G G G 655 A/C G G G 6110 G(GGA) G110delGA G110delGA G110delGA 6110 G(GGA) G110delGA G110delGA G110delGA 0111 D Del Del Del Del $Y112$ Y Del Del Del Sti13 $S(TCC)$ S(TCT) S(TCT) S(TCT) S(TCT) $Y12$ Y Del Del Del Sti13 $S(TCC)$ S(TCT) CCGCT C C Sto2 $85-890$ CCGCT					
$\begin{array}{cccccccc} nt \ 620 & A & G & G & G \\ 624 & G & T & T & T \\ 629/630 & C/A & G/G & G/G & G/G \\ 620 & A & G & G & G \\ 655 & A/C & G & G & G \\ 5108 & S(TCC) & S(TCG) & S(TCG) & S(TCG) \\ 6110 & G(GGA) & G110delGA & G110delGA & G110delGA \\ D111 & D & Del & Del & Del \\ 7112 & Y & Del & Del & Del \\ 5113 & S(TCC) & S(TCT) & S(TCT) & S(TCT) \\ 172 & I & N & I & O \\ 885-890 & CCGCT & CCGCT & CCCGCT^c & CCCGCT \\ 1172 & I & N & I & N \\ 11104 & C & A & C & C \\ 1111 & T & C & A & C & C \\ 1111 & T & C & A & C & C \\ 1111 & T & CGAG & D(GAC) & E(GAG)^d \\ \end{array}$					
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
$\begin{array}{llllllllllllllllllllllllllllllllllll$					-
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$					-
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		· · · ·			
$\begin{array}{cccccc} Y112 & Y & Del & Del & Del & Del \\ S113 & S(TCC) & S(TCT) & S(TCT) & S(TCT) \\ nt 860 & C & C & T^b & C \\ 885-890 & CCGCT & CCGCT & CCCGCT^c & CCCGCT \\ I172 & I & N & I & N \\ nt 1104 & C & A & C & C \\ 1111 & T & C & T & C \\ D183 & D(GAC) & E(GAG) & D(GAC) & E(GAG)^d \end{array}$					
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
nt 860CCTbC885-890CCGCTCCGCTCCCGCTcCCCGCT1172ININnt 1104CACC1111TCTCD183D(GAC)E(GAG)D(GAC)E(GAG)^d					
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
$\begin{array}{ccccccc} nt 1104 & C & A & C & C \\ 1111 & T & C & T & C \\ D183 & D(GAC) & E(GAG) & D(GAC) & E(GAG)^d \end{array}$					
$\begin{array}{cccccccccccccccccccccccccccccccccccc$					
D183 $D(GAC)$ $E(GAG)$ $D(GAC)$ $E(GAG)^d$					
		-			
	D2J4		D(UAC)		D(GAT)

Table 2. Sequence analysis of the chimeric CYP21P/CYP21 gene from CAH patients

CAH, Congenital adrenal hyperplasia; nt, nucleotide; aa, amino acid; del, deletion ^aNucleotide sequence is based on Higashi et al. (1986)

^bT base and C inserted^c are present in the normal CYP21 gene (White et al. 1986)

^dC base changing to G in the case of CH-2 may be a breakpoint site

gested that sequences such as chi-like GCTGGGG (Amor et al. 1988) and the tandem-repetitive minisatellite consensus, GGGCAGGAXG (Jeffreys et al. 1985), may be the recognition sites for the recombinase complex and are responsible for gene recombination. It has been pointed out that the chi-like sequence GCTGGGGG, which is present six times in the *CYP21P* and *CYP21* genes in exons 3–6, may play a role in gene conversion events (Donohoue et al. 1989). Two distinct areas of nt 470 to 999 and nt 1375 to 1993 of the *CYP21* genes have been proposed to be related to crossover regions (Donohoue et al. 1989). Interestingly, chilike sequence GCTGGGC has been found, beginning with the first 5 bp of intron 3 at nt 828 to 834 in CH-1 (Fig. 1B) with no mutation at I172N (nt 999) in exon 4. Therefore, we suggest that the transition for recombination may be anchored in the area between nt 830 to 998. On the other

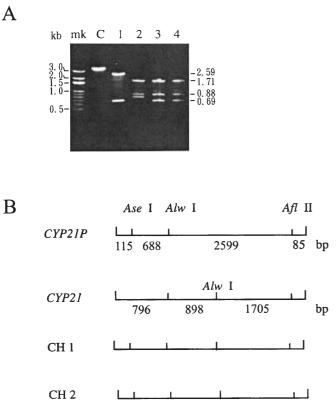


Fig. 2A,B. Analysis of the polymerase chain reaction (PCR) product of 3.5-kb fragments amplified with primers AF1/21BR by restriction endonucleases. A Analysis on 1.5% agarose gel. *Lane C* was a 3.5-kb PCR product of *CYP21P* gene without the digestion of restriction endonucleases. *Lanes 1* and 2 were 3.5-kb fragments of the *CYP21P* and *CYP21* genes, respectively, digested with *Af*II, *Alw*I, and *AseI*. *Lanes 3* and 4 were 3.5-kb fragments of CH-1 and CH-2 clones, respectively, digested with *Af*III, *Alw*I, and *AseI*. *mk* was a 100-bp marker. **B** Schematic representation of restriction cleavage map of *Af*III, *Alw*I, and *AseI* in the *CYP21* genes and CH-1 and CH-2 clones. The location of each restriction site in the *CYP21* genes and CH-1 and CH-2 clones is stated in Results. The size of the *CYP21P* and *CYP21* genes amplified with primers AF1/21BR is 3487 and 3484bp, respectively (Higashi et al. 1986)

hand, sequences that are 90% homologous to a human minisatellite repeat sequence, GGGCAGGAXG, exist in three locations in the CYP21 genes: exon 5 (nt 1117), exon 7 (nt 1958), and a 3' untranslated region (from nt 2710) (Urabe et al. 1990). There has been a report suggesting unequal crossover at sequences in exon 7 (Chu et al. 1992). Surprisingly, in the case of CH-2, a sequence similar to TGGCAGGAGG, as GGGCAGGAXG, is located in D183E (nt 1117) in exon 5 (CH-2, Fig. 1B), which is one of the proposed transition locations. Therefore, we suggest that the base change from C to G (nt 1121) in D183E is evidence for a breakpoint of unequal crossover in this case. We conclude that both minisatellite consensus and chi-like sequences are responsible for initiating interactions between regions to produce these two chimeric genes. From DNA sequencing analysis for the 5' end of these two chimera (Table 2), single base changes, including C instead of T at nt -198 for both CH-1 and CH-2, C for CH-1 and T for CH-2 at nt 395, A for CH-1 and G for CH-2 at nt 502, T insertion at nt 860 in CH-1 and C for both CH-1 and CH-2 at nt 1104, may be polymorphism sites in the *CYP21* genes proposed by our study.

Because the chi-like and minisatellite consensus sequences appear in many loci in the *CYP21* genes, the occurrence of unequal crossover events would be essential.

Acknowledgments This work was supported by King Car Research Foundation from King Car Food Industrial Co., Ltd., Taiwan, The Republic of China.

References

- Amor M, Parker KL, Globerman H, New MI, White PC (1988) Mutation in the *CYP21B* gene (Ile172→Asn) causes steroid 21hydroxylase deficiency. Proc Natl Acad Sci USA 85:1600–1604
- Carroll MC, Palsdottir A, Belt KT, Porter RR (1985) Deletion of complement C4 and steroid 21-hydroxylase genes in the class III region. EMBO J 4:2547–2552
- Chu X, Braun-Heimer L, Rittner C, Schneider PM (1992) Identification of the recombination site within the steroid 21hydroxylase gene (*CYP21*) of the HLA-B47, DR7 haplotype. Exp Clin Immunogenet 9:80–85
- Donohoue PA, Van Dop C, Mclean RH, White PC, Jospe N, Migeon CJ (1986) Gene conversion in salt-losing congenital adrenal hyperplasia with absent complement C4B protein. J Clin Endocrinol Metab 62:995–1002
- Donohoue PA, Jospe N, Migeon CJ, Van Dop C (1989) Two distinct areas of unequal crossingover with the steroid 21-hydroxylase genes produce absence of *CYP21B*. Genomics 5:397–406
- Higashi Y, Yoshioka H, Yamane M, Gotoh O, Fujii-Kuriyama Y (1986) Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: a pseudogene and genuine gene. Proc Natl Acad Sci USA 83:2841–2845
- Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable "minisatellite" regions in human DNA. Nature 314:67–73
- L'Allemand D, Tardy V, Gruters A, Schnabel D, Krude H, Morel Y (2000) How a patient homozygous for a 30-kb deletion of the C4-CYP21 genomoic region can have a nonclassic form of 21hydroxylase deficiency. J Clin Endocrinol Metab 85:4562–4567
- Lee HH (2001) *CYP21* mutations and congenital adrenal hyperplasia. Clin Genet 59:293–301
- Lee HH, Chao HT, Ng HT, Choo KB (1996) Direct molecular diagnosis of CYP21 mutations in congenital adrenal hyperplasia. J Med Genet 33:371–375
- Lee HH, Chao HT, Lee YJ, Shu SG, Chao MC, Kuo JM, Chung Bc (1998) Identification of four novel mutations in the CYP21 gene in congenital adrenal hyperplasia in the Chinese. Hum Genet 103:304–310
- Lee HH, Chang JG, Tsai CH, Tsai FJ, Chao HT, Chung Bc (2000a) Analysis of the chimeric *CYP21P/CYP21* gene in steroid 21hydroxylase deficiency. Clin Chem 46:606–611
- Lee HH, Kuo JM, Chao HT, Lee YJ, Chang JG, Tsai CH, Chung Bc (2000b) Carrier analysis and prenatal diagnosis of congenital adrenal hyperplasia due to 21-hydroxylase deficiency in Chinese. J Clin Endocrinol Metab 85:597–600
- Matteson KJ, Phillips JA III, Miller WL, Chung BC, Orlando PJ, Frisch H, Ferrandez A, Burr IM (1987) P450XXI (steroid 21hydroxylase) gene deletions are not found in family studies of congenital adrenal hyperplasia. Proc Natl Acad Sci USA 84:5858–5862
- Miller WL (1988) Gene conversions, deletions and polymorphisms in congenital adrenal hyperplasia. Am J Hum Genet 42:4–7
- Miller WL, Morel Y (1989) The molecular genetics of 21-hydroxylase deficiency. Annu Rev Genet 23:371–393
- Rodrigues NR, Dunham I, Yu CY, Carroll MC, Porter RR, Campbell RD (1987) Molecular characterization of the HLA-linked steroid 21-hydroxylase B gene from an individual with congenital adrenal hyperplasia. EMBO J 6:1653–1661

- Speiser PW, Agedere L, Ueshiba H, White PC, New MI (1991) Aldosterone synthesis in salt-wasting congenital adrenal hyperplasia with complete absence of adrenal 21-hydroxylase. N Engl J Med 324:145– 149
- Tusie-Luna MT, White PC (1995) Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase gene) and CYP21P involve different mechanisms. Proc Natl Acad Sci USA 92:10796–10800
- Urabe K, Kimura A, Harada F, Iwanaga T, Sasazuki T (1990) Gene conversion in steroid 21-hydroxylase genes. Am J Hum Genet 46:1178–1186
- White PC, New MI, Dupont B (1984) HLA-linked congenital adrenal hyperplasia results from a defective gene encoding a cytochrome P-450 specific for steroid 21-hydroxylation. Proc Natl Acad Sci USA 81:7505–7509
- White PC, New MI, Doupont B (1986) Structure of human steroid 21hydroxylase genes. Proc Natl Acad Sci USA 83:5111–5115
- White PC, New MI, Dupont B (1987) Congenital adrenal hyperplasia. N Engl J Med 316:1519–1524, 1580–1586
- White PC, Vitek A, Dupont B, New MI (1988) Characterization of frequent deletions causing steroid 21-hydroxylase deficiency. Proc Natl Acad Sci USA 85:4436–4440