ABSTRACT Four steroid 21-hydroxylase B [P-450(C21)B] genes (designated P.7, P.10-1, P.10-2, and P.3) from three P-450(C21)-deficient patients were isolated to analyze their structures and functions. Several base changes were observed in the sequences of the four P-450(C21)B genes as compared to that of the functional B gene. Many of these base changes were identical to those of the P-450(C21)A pseudogene. The three DNAs (P.10-1, P.10-2, and P.3) produced no P-450(C21) activity in a functional assay for P-450(C21) by the COS cell expression system, while the P.7 DNA expressed the activity. The P.10-1 and P.10-2 DNAs were shown to have a point mutation in the second intron, causing aberrant splicing. The P.3 DNA carried three clustered missense mutations in the sixth exon, which impaired P-450(C21) activity. All these critical mutations could be seen in the corresponding site of the P-450(C21)A pseudogene. These data strongly suggest the involvement of gene conversion in this genetic disease.

Steroid 21-hydroxylase [P-450(C21)], steroid 21-monoxygenase; steroid, hydrogen-donor:oxygen oxidoreductase (21-hydroxylating), EC 1.14.99.10] deficiency (1, 2) is the most frequent cause (>95%) of congenital adrenal hyperplasia, one of the most common inborn errors of metabolism, affecting 1 in 15,000 births. This disease is inherited as a monogenic autosomal recessive trait closely linked to HLA major histocompatibility complex (3). P-450(C21) is involved in the conversion of progesterone and 17-hydroxyprogesterone to 11-deoxycorticosterone and 11-deoxy cortisol, the intermediate steps in mineralocorticoid and glucocorticoid biosynthesis, respectively (4). By the defect of the P-450(C21) gene, synthesis of aldosterone and/or cortisol is impaired and 17-hydroxyprogesterone consequently accumulates, leading to excess production of androgen. These hormonal conditions develop three types of clinical symptoms, referred to as the salt-wasting, simple-virilizing, and nonclassical forms (5).

Recent gene cloning analysis showed that there are two P-450(C21) genes—P-450(C21)A and P-450(C21)B—each of which is located immediately adjacent to the 3' end of the two genes encoding the fourth component of complement, C4A and C4B, which reside in the class III region of the HLA complex (6, 7). The B gene characterized by the 3.7-kilobase (kb) TaqI fragment of the genomic DNA is functional, whereas the A gene of the 3.2-kb TaqI fragment is a pseudogene harboring at least three deleterious mutations in the coding sequence (8, 9). Nonetheless, these two genes are extremely homologous (98%) even in the introns and the 5' flanking sequence, suggesting their concerted evolution by exchanging genetic information with each other (8). Genomic blot analysis of P-450(C21)-deficient patients by using cloned DNAs as probes (10–13) revealed that some patients had the B gene either totally deleted or converted to the A gene. Further analysis with oligonucleotide probes specific for the three deleterious sequences in the A gene could show that at least one allele of the B gene (3.7-kb TaqI fragment) in a patient was inactivated by being replaced in part by a harmful mutation in the A gene (14). In most cases, however, patients appear to have the B genes in their genomes, judging from the size of the TaqI fragment, suggesting that some other small mutations, which cannot be detected by blot analysis, may possibly be involved in inactivation of the B gene. To clarify how such small mutations cause P-450(C21) deficiency, we isolated several B genes from patients and characterized their structures and functions. In the present paper, we provide findings that an altered splicing mechanism and amino acid substitutions are at least in part the cause of P-450(C21) deficiency.

MATERIALS AND METHODS

Patients. Patients with P-450(C21) deficiency were followed by the Division of Pediatric Endocrinology at National Children's Hospital. Their clinical diagnoses and HLA haplotypes have been reported (14).

Cloning and Sequence Analysis. Genomic DNA samples from patients were prepared from peripheral blood lymphocytes transformed by Epstein–Barr virus as described (14). Genomic DNA libraries were constructed from 15-kb BamHI fragments, including the P-450(C21)B gene and BamHI-digested Charon 28 vector. The libraries were screened first with the entire sequence of the P-450(C21) gene and then with the B-gene-specific oligonucleotide probe (2-B probe in ref. 14). The isolated P-450(C21)B genes were characterized by restriction mapping and DNA blot analysis and finally sequenced by the dideoxy nucleotide method (15).

Expression of the Isolated P-450(C21) Genes. The 2.8-kb AccI/KpnI (Asp-718) fragments, which cover the sequence of the P-450(C21) gene from –75 to +2730, were isolated and filled up by the Klenow fragment. These blunt-ended fragments were inserted into a HincII site in the polylinker sequence of pUC18. The genomic DNA fragments were regenerated by digestion of those plasmids with BamHI and HindIII and inserted between a BamII and a HindIII site of pKCRH-2 (16), an expression vector in COS cells. The resultant plasmid was transfected into COS-7 cells by the DEAE-dextran (17) or calcium phosphate coprecipitation (18) method. After 3 days of incubation in Dulbecco's
modified Eagle's medium containing 15% fetal calf serum, [14C]progesterone or 17-hydroxy[14C]progesterone (1–2 μM) was added to the culture medium and the cells were incubated for a further 24 hr. Steroids in the culture fluids were extracted with dichloromethane and chromatographed on chloroform/acetone (70:30, vol/vol). The separated products on the plate were subjected to autoradiography. Steroid products were identified by comigration with authentic standards (19).

Nuclease S1 Mapping Analysis. Nuclease S1 mapping analyses were performed with 50 μg of the total RNA and 32P-labeled BstXI/Nco I fragment of ~220 base pairs (bp) (≥2 × 106 cpm/μg) as described (20).

RESULTS
Isolation of the P-450(C21)B Genes from Patients and Sequence Analysis. Three patients, of which two are salt wasting and one is simple virilizing (patient nos. 3, 7, and 10 in ref. 14), were selected for DNA analysis because they have the B gene (3.7-kb Taq I fragment) in their genomes with no apparent defect on DNA blot analysis (14). Their genomic DNAs were subjected to gene cloning and clones for one B gene each from patient nos. 7 and 3 and two from patient no. 10 were finally isolated. The two genes from patient no. 10 turned out to differ from each other; that is, they were from different alleles because the two 15-kb BamHI fragments showed polymorphism in the Taq I restriction site (data not shown) and had a different sequence in the third and seventh exons (Fig. 1). The four isolated B genes were sequenced from the Acc I site (~75) to the Kpn I site (~2730) and the results are summarized in Fig. 1.

Neither frameshift nor nonsense mutation, which can apparently prevent synthesis of a functional protein, could be detected in this sequence analysis. Several base changes, however, were observed in all four B genes in reference to the sequence of the functional B gene that we originally isolated (8). Some of those base changes, which accompany amino acid substitutions, may possibly be fatal to the function of P-450(C21), or those in the intronic sequence should affect normal RNA processing. To address these possibilities, we tried to express the cloned B genes in COS cells for examination of P-450(C21) activity.

Expression of the Cloned B Genes in COS Cells. Zuber et al. (19) and O. Minowa, K. Sogawa, and Y.F.K. (unpublished observation) have recently described the functional expression of microsomal P-450 cDNAs in COS cells. For analysis of the function of the genes, we took advantage of this expression system except that we used the cloned genomic DNAs instead of the cDNAs.

mRNAs for P-450(C21) were detected in the COS cells transfected with the DNAs of the patients' B genes or the functional B gene (Fig. 2a, lanes 1–4 and 6). The mRNA derived from the A gene was not detected at the correct size (lane 5), although the faint smeared band of somewhat larger size was observed. The reason for the absence of A gene

![Diagram](image-url)
transcripts is not clear, but this observation is in agreement with the fact that the transcripts from the A gene were not detected in the human adrenal RNA (21).

No P-450(C21) activity could be detected in the COS cells transfected with P.10-1, P.10-2, P.3, or A (negative control) DNAs, while the P.7 and B (positive control) DNAs produced the P-450(C21) activities to a similar level (Fig. 2b). The differential activities between the two substrates in each transfected cell could not be observed in our assay system. The three inactive B genes from two patients (P.10 and P.3) must have defective mutations within the 2.8-kb genomic fragment. These mutations should be ascribed to one or several of the nucleotide alterations found in DNA from each patient.

Identification of the Defective Mutations. We could localize the deleterious mutations in the three patients' DNAs by constructing a series of hybrid genes between the active (e.g., B or P.7) and the inactive (P.10-1, P.10-2, and P.3) DNAs and subsequently testing for their P-450(C21) activities expressed in the COS cells. A similar strategy was undertaken to determine the oncogenic region of the activated ras gene product (22). After repeating such experiments with these hybrid genes, we could finally determine the causative mutations to the inactivation of the three patient genes.

As shown in Fig. 3, both P.10-1 and P.10-2 DNAs carry an identical defective mutation in the second intron, whereas the P.3 DNA contains another in the sixth exon. A 42-bp Pvu II/PstI fragment containing one base alteration in the second intron (from C to G at +656) of the P.10-1 and P.10-2 DNAs was replaced with that of the functional B gene or vice versa to examine its ability to produce the P-450(C21) activity. The results are shown in Fig. 3a. Similar experiments were performed with the P.3 DNA using its 90-bp Bsm I/Pvu II fragment in the sixth exon containing the clustered three missense mutations from T to A at 1380, 1383, and 1389, respectively (Fig. 3b).

Since the defective point mutations identified in the second intron of the P.10-1 and P.10-2 DNAs change the sequence from AC to AG, which is identical to the consensus sequence of the splicing junction and since the P.10-1 and P.10-2 DNAs produced mRNA of apparently similar amount and size to those of the functional B gene on RNA blot analysis (Fig. 2a), we assume that the mutation should disturb the normal splicing without the gross change in mRNA synthesis; for example, the splicing site might be shifted to the nucleotide just after the G mutation in the intron, which results in synthesis of an immature or defective protein.

To examine whether it really occurs, we performed nuclease S1 mapping analysis with the BstXI/Neo I DNA fragment used as a probe (Fig. 3a). The results are shown in Fig. 4. The expected size of the protected fragment produced by normal splicing is detected in the samples from the B and P.7 DNAs (50, lanes 1 and 2). However, no band could be observed at the corresponding position for P.10-1 (lane 3). Alternatively, we could find several protected bands of larger sizes (S1-S3). The major one (S1) corresponds to a position 19 bp upstream from the normal splicing site, or 7 bp further upstream than we had expected if the mutated G together with the preceding A functioned as a splicing acceptor signal. This result indicates that another upstream AG sequence should be capable of functioning as a splicing signal. The altered splicing at this site causes a frameshift in the following third exon to generate the termination codon at Lys-103 in the third exon, resulting in production of the truncated protein. Such a small protein must be inactive because the heme-binding site essential for P-450 activity is supposed to be in the 10th exon (23). Other minor bands (S2 or S3) seem to be derived from the splicing at the further upstream consensus (AG) or consensus-like (AC) sequence as shown in Fig. 4. These aberrant splicing products should cause a frameshift mutation in the third exon (S2) or the unusual protein with an insertion of 11 extra amino acid residues (S3). Neither of them can be functional because we could not detect P-450(C21) activity in the COS cells transfected with the P.10-1 DNA. Sequence analysis revealed that the P.7 DNA (lane 2) carried an A in place of a G as shown in Fig. 1, but it still produced the normal splicing transcript (lane 2). These results clearly show that a point mutation from C (or A) to G near the splice junction completely abolishes normal splicing and generates aberrant splicing junctions. Unexpectedly, the cells transfected with the functional B or the P.7 DNA appeared to produce also aberrantly spliced mRNAs of similar sizes to those of the P.10-1 DNA (Fig. 4), although in a minor amount. This is discussed below.

DISCUSSION

The P.7 DNA appeared to express normally the P-450(C21) activity in our COS cell system, although it was isolated from the patient. Considering that the genomic sequence used for construction of the expression plasmid covers the region from -75 to +2730, it is concluded that no defective mutation exists, at least in this genomic region, mostly consisting of the structural gene of the P-450(C21) gene. It is
plausible, therefore, to infer that the sequences outside of this region, which may play a regulatory role for gene expression, carry harmful mutations. This inference is now amenable to the experimental analysis of reverse genetics and should be tested soon.

The clustered three missense mutations located in the sixth exon of P.3 DNA change isoleucine, valine, and methionine to asparagine, glutamic acid, and lysine, respectively. We do not yet know which or how many of the three missense mutations are involved in the inactivation of the B gene. Indeed, some of these mutations may be deleterious, since this region is included in a putative substrate binding portion, as suggested by the comparative analysis of all the known primary structures of P-450s (23).

The presence of the aberrantly spliced mRNA in the COS cells transfected with the functional B or P.7 DNAs has been demonstrated by nuclease S1 mapping analysis (Fig. 4), although in a small amount. Although these observations might be due in part to an artifact from forced expression under the simian virus 40 early promoter in COS cells, these phenomena were also observed by RNA blot analysis of the human adrenal mRNA; the weak, but still detectable, signal could be observed with the oligonucleotide probes specific for the intronic sequence preceding the junction between the second intron and the third exon (data not shown). It is possible, therefore, that the truncated small protein or unusual protein with the inserted residues can be produced in the adrenal tissue of normal individuals. We do not know, however, whether these unusual transcripts have any biological significance. The G point mutation in the second intron of the B gene from patient no. 10 completely abolished the normal splicing with a concomitant increase in the usage of the aberrant one. The rationale for this switching mechanism has not been clarified, but it will be interesting to study for better understanding of the splicing mechanism in general.

We were interested in determining what population of patients carry the mutations that we have clarified in this paper. For this analysis, we prepared the oligonucleotide probes specific for those mutations as well as for the functional counterparts and surveyed the 11 patients by DNA blot analysis. The results indicate that the G mutation seems to be relatively common in patients' genomes (unpublished results). Furthermore, a 3.7-kb gene reported recently by Rodrigues et al. (24) also has this mutation. In addition, the Nco I polymorphism found in the 3.7-kb gene could also be observed in our patients' DNAs; both P.10-1 and P.10-2 have the additional Nco I site in the seventh exon (C at +1645, as shown in Fig. 1).

At the same site of the G point mutation in the second intron, another base change of A was observed in the P.7 and P.3 DNAs. The A mutation is silent since the product from the P.7 DNA was active in our assay and the defect in the P.3 gene was mapped in the other region (sixth exon). The functional B gene we originally identified has a C at this site.

Biochemistry: Higashi et al.


FIG. 3. Identification of defective sequence in patients' B genes. Replacement experiments of P.10-1 and -2 (a) or P.3 (b) with the functional B gene are schematically shown along with the corresponding region of their sequences. The sequence is given in full for the B gene and only altered nucleotides are described for P.10-1 (a) and P.3 (b) under the B gene sequence. Open boxes represent the B gene sequence. The P.10-1, P.10-2, and P.3 DNA sequences are indicated by shaded, stippled, and hatched boxes, respectively.
Fig. 4. Examination of aberrant splicing site by nuclease S1 mapping analysis. Fifty micrograms of total RNA, each derived from COS cells transfected with the B (lane 1), P (lane 2), and P.10-1 (lane 3) DNA and 32P-labeled Nco I/BstXI probe (see Fig. 3a) was used for nuclease S1 mapping analysis. To visualize the ladder of Maxam–Gilbert more clearly (lane AG/CT), the longer exposure of autoradiograms was superimposed on the shorter one. S0, protected band derived from the normal splicing site; S1–S3, protected bands from aberrant splicing sites. Corresponding genomic sequence is shown on the left. The possible consensus AG sequences are underlined.

Southern blot analysis of 11 patients using the oligonucleotide probes, each specific for one of the three sequences, revealed that the A sequence occurs as frequently as the C sequence (data not shown). Thus, this position is highly polymorphic.

One of the most important observations in the present study, as well as in a previous one (14), is that the 3.2-kb A and 3.7-kb B alleles of P-450(C21) are highly polymorphic. All four DNAs we characterized have different sequences, and our Southern blot analysis using site-specific oligonucleotide probes has so far identified the several different alleles for the 3.2-kb and 3.7-kb gene including the defective mutant alleles. These mutant alleles have been identified so far, one containing an 8-bp deletion in the second exon identified in patient no. 9 in a previous report (14) and two others identified in this report. The causative mutations in the 3.7-kb B genes of other patients remain to be seen.

Frequent occurrence of gene conversion in the P-450(C21) genes has been suggested (8) and recently demonstrated by our group (14) and by others (13, 25). Interestingly, more than half of the base changes identified in the four isolated B genes are identical to those of the A pseudogene, and those base changes are not restricted to a certain region but rather are scattered throughout the entire sequence of the P-450(C21) gene (Fig. 1). These observations can be best explained by a model of the molecular mechanism for gene conversion, which is supposed to result from mismatch repair of a heteroduplex between the nearly homologous genes (26). Thus, gene conversion events have frequently taken place in the P-450(C21) gene, which should cause high allelic polymorphism and, as a result, a high incidence of P-450(C21) deficiency.

Note. Gonzalez et al. (27) have recently published a paper suggesting that aberrant splicing mechanisms in P-450b1 cause deficient metabolism of a group of drugs in humans.

We thank Drs. Mishina and Numa for providing us with the expression vector pKCRH-2. This work was partly supported by a grant from the Ministry of Health and Welfare, Japan, by a grant from the Ministry of Education, Science and culture, Japan, and by funds obtained from the Watanabe Memorial Foundation for the Advance ment of Science and Technology and under the Life Science Project from the Institute of Physical and Chemical Research, Japan.