Congenital Adrenal Hyperplasia due to 21-Hydroxylase Deficiency*

PERRIN C. WHITE AND PHYLLIS W. SPEISER

Division of Pediatric Endocrinology (P.C.W.), University of Texas Southwestern Medical Center, Dallas, Texas 75390-9063; and Division of Pediatric Endocrinology (P.W.S.), North Shore University Hospital and New York University School of Medicine, Manhasset, New York 11030

ABSTRACT

More than 90% of cases of congenital adrenal hyperplasia (CAH, the inherited inability to synthesize cortisol) are caused by 21-hydroxylase deficiency. Females with severe, classic 21-hydroxylase deficiency are exposed to excess androgens prenatally and are born with virilized external genitalia. Most patients cannot synthesize sufficient aldosterone to maintain sodium balance and may develop potentially fatal “salt wasting” crises if not treated. The disease is caused by mutations in the CYP21 gene encoding the steroid 21-hydroxylase enzyme. More than 90% of these mutations result from intergenic recombinations between CYP21 and the closely linked CYP21P pseudogene. Approximately 20% are gene deletions due to unequal crossing over during meiosis, whereas the remainder are gene conversions—transfers to CYP21 of deleterious mutations normally present in CYP21P. The degree to which each mutation compromises enzymatic activity is strongly correlated with the clinical severity of the disease in patients carrying it. Prenatal diagnosis by direct mutation detection permits prenatal treatment of affected females to minimize genital virilization. Neonatal screening by hormonal methods identifies affected children before salt wasting crises develop, reducing mortality from this condition. Glucocorticoid and mineralocorticoid replacement are the mainstays of treatment, but more rational dosing and additional therapies are being developed. (Endocrine Reviews 21: 245–291, 2000)

I. Introduction

II. Biochemistry of CAH
   A. Biochemistry of normal steroid synthesis
   B. Regulation of adrenal steroid secretion
   C. Abnormal steroids in 21-hydroxylase deficiency

III. Pathophysiology of CAH
   A. Normal sexual differentiation
   B. Normal prenatal development of adrenal glands
   C. Adrenarche
   D. Prenatal virilization
   E. Salt wasting
   F. Postnatal signs of androgen excess
   G. Reproductive function in classic CAH
   H. Neuropsychology of CAH
   I. Tumors
   J. Nonclassic CAH phenotypes
   K. Heterozygotes

IV. Diagnosis of 21-Hydroxylase Deficiency
   A. Evaluation of ambiguous genitalia
   B. Newborn screening
   C. Further biochemical evaluation

V. Treatment
   A. Glucocorticoid replacement
   B. Mineralocorticoid replacement
   C. Other therapeutic approaches
   D. Corrective surgery
   E. Psychological counseling

F. Treatment of precocious puberty
G. Prenatal therapy

VI. Molecular Genetic Analysis
   A. Biochemistry of CYP21
   B. Structure-function relationships
   C. CYP21 gene structure
   D. Transcription
   E. HLA linkage
   F. Mutations causing 21-hydroxylase deficiency
   G. De novo recombinations
   H. Mutation detection and approaches to prenatal diagnosis
   I. Correlations between genotype and phenotype
J. Why is CAH so common?

VII. Summary

Virilizing congenital adrenal hyperplasia (CAH) is the most common cause of genital ambiguity, and 90–95% of CAH cases are caused by 21-hydroxylase deficiency. Females affected with severe, classic 21-hydroxylase deficiency are exposed to excess androgens prenatally and are born with virilized external genitalia. First described in the mid-19th century, a more thorough understanding of this disease was not forthcoming until the mid-20th century, when the recessive nature of the genetic trait and identification of hormonal abnormalities were recognized (1).

The fundamental defect among patients with CAH due to 21-hydroxylase deficiency is that they cannot adequately synthesize cortisol. Inefficient cortisol synthesis signals the hypothalamus and pituitary to increase CRH and ACTH, respectively. Consequently, the adrenal glands become hyperplastic. But rather than cortisol, the adrenals produce excess sex...
hormone precursors that do not require 21-hydroxylation for their synthesis. Once secreted, these hormones are further metabolized to active androgens—testosterone and dihydrotestosterone—and to a lesser extent estrogens—estrone and estradiol. The net effect is prenatal virilization of girls and rapid somatic growth with early epiphyseal fusion in both sexes. About three-quarters of patients cannot synthesize sufficient aldosterone to maintain sodium balance and are termed “salt wasters.” This predisposes them to episodically develop potentially life-threatening hyponatremic dehydration.

Patients with sufficient aldosterone production and no salt wasting who have signs of prenatal virilization and/or markedly increased production of hormonal precursors of 21-hydroxylase (e.g., 17-hydroxyprogesterone), are termed “simple virilizers.” In addition, a mild nonclassic form of the disorder is recognized in which affected females have little or no virilization at birth (Table 1).

It has now been 15 yr since the CYP21 gene encoding the steroid 21-hydroxylase enzyme was demonstrated to be affected in patients with 21-hydroxylase deficiency (2), and it seemed an appropriate time to comprehensively review subsequent progress in understanding this disorder. References to earlier work can be found in previous reviews (1, 3–5).

II. Biochemistry of CAH

A. Biochemistry of normal steroid synthesis

The rate-limiting step in steroid biosynthesis is importation of cholesterol from cellular stores to the matrix side of the mitochondria-inner membrane where the cholesterol side chain cleavage system (CYP11A, adrenodoxin, adreno- doxin reductase) is located. This is controlled by the steroidogenic acute regulatory (StAR) protein (6), the synthesis of which is increased within minutes by trophic stimuli such as ACTH or, in the zona glomerulosa, increased intracellular calcium. StAR is a synthesized as a 37-kDa phosphoprotein that contains a mitochondrial importation signal peptide. However, importation into mitochondria is not necessary for StAR to stimulate steroidogenesis, and it now seems likely that, to the contrary, mitochondrial importation rapidly inactivates StAR (7). The mechanism by which StAR mediates cholesterol transfer across the mitochondrial membrane is not yet known.

It is clear that StAR is not the only protein that mediates cholesterol transfer across the mitochondrial membrane. Another protein that appears necessary (but not sufficient, at least in the adrenals and gonads) for this process is the so-called peripheral benzodiazepine receptor, an 18-kDa protein in the mitochondrial outer membrane that is complexed with the mitochondrial voltage-dependent anion carrier in contact sites between the outer and inner mitochondrial membranes (8). This protein does not appear to be directly regulated by typical trophic stimuli, but it is stimulated by endozepines, peptide hormones also called diazepam-binding inhibitors. Endozepines may be regulated by ACTH to some extent, but not with a rapid time course. Thus far, it is not yet clear whether there is a direct physical

<table>
<thead>
<tr>
<th>Phenotype:</th>
<th>Classic salt wasting</th>
<th>Classic simple virilizing</th>
<th>Nonclassic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis</td>
<td>Newborn-6m</td>
<td>2–4 y</td>
<td>Child-adult</td>
</tr>
<tr>
<td>Genitalia</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Aldosterone</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Renin</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Cortisol</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>17-OH-progesterone</td>
<td>&gt;20,000 ng/dl</td>
<td>&gt;10,000–20,000 ng/dl</td>
<td>1,500–10,000 ng/dl (ACTH-stimulated)</td>
</tr>
<tr>
<td>Testosterone</td>
<td>↑ In pre-puberty only</td>
<td>↑ In pre-puberty only</td>
<td>Variably ↑ in pre-puberty only</td>
</tr>
<tr>
<td>Treatment</td>
<td>Glucocorticoid + mineralocorticoid (+ sodium)</td>
<td>Glucocorticoid (+ mineralocorticoid)</td>
<td>Glucocorticoid, if symptomatic</td>
</tr>
<tr>
<td>Somatic growth*</td>
<td>-2–3 SD, husky-obese</td>
<td>-1–2 SD</td>
<td>-1 SD</td>
</tr>
<tr>
<td>Incidence*</td>
<td>1/20,000</td>
<td>1/10,000</td>
<td>1/1000</td>
</tr>
<tr>
<td>Typical mutations*</td>
<td>Deletion</td>
<td>I172N</td>
<td>V281L</td>
</tr>
<tr>
<td></td>
<td>nt 656g (&quot;intron 2 g&quot;)</td>
<td>nt 656g</td>
<td>P30L</td>
</tr>
<tr>
<td>I236N/V237E/M239K</td>
<td>G110D</td>
<td>Q318X</td>
<td></td>
</tr>
<tr>
<td>Q318X</td>
<td>R356W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Enzymatic activity*</td>
<td>0</td>
<td>1</td>
<td>20–50</td>
</tr>
</tbody>
</table>

* SD, Standard deviation scores.
* Incidence in general white population. See Table 3 for incidence of classic disease (salt wasting plus simple virilizing) from neonatal screening in various populations.
* See Table 4 and Section VI.F.
* Enzymatic activity predicted from in vitro expression studies (see Section VI.F).
interaction between StAR and the peripheral benzodiazepine receptor.

The first enzymatic step in steroid synthesis (Fig. 1) is the conversion of cholesterol, a C27 compound, to the C21 steroid pregnenolone (reviewed in Ref. 9). This is catalyzed by the mitochondrial cytochrome P450 enzyme CYP11A (P450 scc, cholesterol desmolase, side-chain cleavage enzyme; see Ref. 10 for further description of the CYP and P450 enzyme terminology). Pregnenolone is the common precursor for all other steroids and, as such, may undergo metabolism by several other enzymes.

To synthesize mineralocorticoids in the zona glomerulosa, 3β-hydroxysteroid dehydrogenase (3β-HSD) in the endoplasmic reticulum and mitochondria (11) converts pregnenolone to progesterone (12). This is 21-hydroxylated in the endoplasmic reticulum by CYP21 (P450c21, 21-hydroxylase) to produce deoxycorticosterone (DOC). Aldosterone, the most potent 17-deoxysteroid with mineralocorticoid activity, is produced by the 11β-hydroxylation of DOC to corticosterone (historically termed compound B), followed by 18-hydroxylation and 18-oxidation of corticosterone. The final three steps in aldosterone synthesis are accomplished by a single mitochondrial P450 enzyme, CYP11B2 (P450aldo, aldosterone synthase, reviewed in Ref. 13).

To produce cortisol, the major glucocorticoid in man, CYP17 (P450c17, 17α-hydroxylase/17, 20 lyase) in the endoplasmic reticulum of the zona fasciculata and zona reticularis converts pregnenolone to 17α-hydroxypregnenolone (14). 3β-Hydroxysteroid dehydrogenase in the zona fasciculata utilizes 17α-hydroxypregnenolone as a substrate, producing 17α-hydroxyprogesterone. The latter is 21-hydroxylated by CYP21 to form 11-deoxycorticisol, which is converted to cortisol by CYP11B1 (P450c11, 11β-hydroxylase) in mitochondria.

In the zona reticularis of the adrenal cortex and in the gonads, the 17,20-lyase activity of CYP17 converts 17α-hydroxyprogrenolone to dehydroepiandrosterone (DHEA, a C19 steroid and sex hormone precursor). DHEA is further converted by 3β-HSD to androstenedione. In the gonads, this is reduced by 17β-hydroxysteroid dehydrogenase to testosterone [there are several isozymes of 17β-hydroxysteroid dehydrogenase, some of which possess both oxidative and reductive activity (15)]. In pubertal ovaries, aromatase (CYP19, P450c19) can convert androstenedione and testosterone to estrone and estradiol, respectively (16). Testosterone may be further metabolized to dihydrotestosterone by steroid 5α-reductase in androgen target tissues (17).

B. Regulation of adrenal steroid secretion

1. Cortisol secretion. Cortisol secretion is regulated mainly by ACTH. ACTH is a 39-amino acid peptide that is produced in
the anterior pituitary. It is synthesized as part of a larger mol wt precursor peptide, POMC. This peptide is also the source of β-lipoprotein (β-LPH). In addition, ACTH and β-LPH are cleaved further to yield α-MSH and β-MSH, γ-LPH, β- and γ-endorphin, and enkephalin. The POMC precursor peptide is found in a variety of extrahypothalamic tissues, including the gastrointestinal tract, numerous tumors, and the testis. It is secreted in small amounts from the anterior pituitary gland and does not bind significantly to the ACTH receptor. Another pro-ACTH fragment, corticotropin-like intermediate lobe peptide (CLIP), is made in the rodent anterior pituitary, but not in the normal human pituitary (18).

ACTH acts through a specific G protein-coupled receptor to increase levels of cAMP (19). cAMP has short-term (minutes to hours) effects on cholesterol transport into mitochondria (6) but longer term (hours to days) effects on transcription of genes encoding the enzymes required to synthesize cortisol (20). The transcriptional effects occur, at least in part, through increased activity of protein kinase A, but it is not known whether the targets of this kinase act directly or indirectly on CYP21 (see Section V.D.3). ACTH also influences the remaining steps in steroidogenesis as well as the uptake of cholesterol from plasma lipoproteins. It also maintains the size of the adrenal glands. In addition to these effects on the adrenal gland, it stimulates melanocytes and results in hyperpigmentation when secreted in excess, as occurs in Addison’s disease.

CRH is the principal hypothalamic factor that stimulates the pituitary production of ACTH (21, 22). Vasopressin, a peptide product of the posterior pituitary gland, also stimulates ACTH release by acting synergistically with CRH and is an important physiological regulator of ACTH (23). CRH is produced in the paraventricular nuclei of the hypothalamus and is also found in other parts of the central nervous system and in other locations such as peripheral leukocytes. Paracrine action of hypothalamic peptides, e.g., vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase-activating polypeptide (PACAP), plays a role in CRH release (24). Hypothalamic CRH is transported to the anterior pituitary cells by the hypophysial portal vessels. CRH activates ACTH secretion via a specific receptor coupled to CAMP-dependent signaling. CRH is secreted in a pulsatile fashion that results in the episodic secretion of ACTH and in the diurnal variation of cortisol secretion. The magnitude of the cortisol response to each ACTH burst remains relatively constant. Therefore, it is the number of secretory periods, rather than the magnitude of each pulse of CRH and ACTH, that determines the total daily cortisol secretion.

Numerous factors, such as metabolic, physical, or emotional stress, influence levels of glucocorticoid secretion, mediated by ACTH secreted in response to hypothalamic secretion of CRH and vasopressin. As noted above, paracrine action of various peptides may contribute to modulation of hormone production in the hypothalamus, pituitary, and adrenal. Cortisol is the primary negative regulator of resting activity of the hypothalamic-pituitary-adrenal (HPA) axis through negative feedback on ACTH and CRH secretion. Furthermore, it may inhibit some of the higher cortical activities that lead to CRH stimulation.

The negative feedback effects of cortisol are exerted at the level of both the hypothalamus and the pituitary and are mediated by Type II corticosteroid receptors (i.e., classic glucocorticoid receptors) (25). Whether and to what extent direct glucocorticoid feedback on the adrenal cortex itself regulates cortisol synthesis is unclear. In vitro studies using rat adrenocortical cells suggest that corticosterone may act to inhibit steroidogenesis (26). Northern blotting demonstrates that glucocorticoid receptors are expressed in human adrenals (27), but a physiological role in direct negative regulation of cortisol secretion has not been demonstrated.

2. Aldosterone secretion. The rate of aldosterone synthesis, which is normally 100- to 1,000-fold less than that of cortisol synthesis, is regulated mainly by angiotensin II and potassium levels, with ACTH having only a short-term effect (28). Angiotensin II occupies a G protein-coupled receptor-activating phospholipase C (29). The latter protein hydrolyzes phosphatidylinositol bisphosphate to produce inositol triphosphate and diacylglycerol, which raise intracellular calcium levels and activate protein kinase C and calmodulin-dependent protein (CaM) kinases. Similarly, increased levels of extracellular potassium depolarize the cell membrane and increase calcium influx through voltage-gated L-type calcium channels (30). Phosphorylation of as yet unidentified factors by CaM kinases increases transcription of the aldosterone synthase (CYP11B2) enzyme required for aldosterone synthesis (28); as yet, the pathways influencing 21-hydroxylase (CYP21) expression in the zona glomerulosa have not been elucidated.

C. Abnormal steroids in 21-hydroxylase deficiency

1. Elevated 17-hydroxyprogesterone. The most characteristic biochemical abnormality in 21-hydroxylase deficiency is elevation of 17-hydroxyprogesterone (17-OHP), the main substrate for the enzyme. Basal serum 17-OHP values usually exceed 10,000 ng/dl, although about 10% of severely affected infants have low initial levels in the newborn period (31), especially if levels are obtained on the first day of life. Differentiation of 21-hydroxylase deficiency from other forms of CAH may be accomplished by both clinical features of the disease (Table 2) and by the complete adrenocortical hormone profile comparing precursor to product ratios after ACTH stimulation. It is important to realize that without a complete adrenocortical profile, other steroidogenic defects—both 3β-HSD (32) and 11β-hydroxylase deficiency (34)—may be misdiagnosed as 21-hydroxylase deficiency. This has significant bearing on medical therapy since 11β-hydroxylase patients are often hypertensive and require specific therapy for this problem. Moreover, these assays should be performed in laboratories with high standards for quality control, including preliminary chromatography, to avoid problems of cross-reactivity when some hormone levels are extremely high (33). This can be a serious concern when the choice of laboratory is limited in a managed care environment.

The highest 17-OHP levels (up to 100,000 ng/dl after ACTH stimulation) are seen in patients with the salt wasting form of the disease. Simple virilizing patients tend to
have somewhat lower levels, although the range overlaps that seen in salt wasting patients (34). The milder, nonclassic form of CAH manifests even less markedly elevated hormone levels, especially in the newborn period. Nonclassic patients are most reliably diagnosed by their response to ACTH stimulation (35); random measurements of basal serum 17-OHP may be normal in mildly affected nonclassic patients unless performed in the early morning (i.e., before 0800 h). Compound heterozygotes for classic and nonclassic CYP21 mutations (see Section VI.F) tend to have somewhat higher ACTH-stimulated 17-OHP levels than individuals homozygous for nonclassic mutations (36). Hormonal testing is not very sensitive for identification of heterozygotes when 17-OHP is used as a marker. In one study, only 50% of obligate heterozygotes had 17-OHP measurements after ACTH stimulation that differed from those of genotypically normal individuals (37). Heterozygotes are more readily identified when one examines the ratio of 17-OHP to cortisol (38).

2. Other abnormal steroids. Other hormones that are elevated in untreated CAH include progesterone, androstenedione, and, to a lesser extent, testosterone. An abnormal steroid, 21-deoxycortisol, is characteristically elevated (39–41). DHEA, the main adrenal 19-carbon steroid product, is not a good marker of 21-hydroxylase activity. DHEA-sulfate (DHEAS) binds with high affinity to albumin, has a long plasma half-life, and as such is not very responsive to acute ACTH stimulation. Diagnostic assays are discussed in Section IV.C.

III. Pathophysiology of CAH

A. Normal sexual differentiation

Early in gestation, the gonads are indifferent and bipotential (Fig. 2). During the 7th week, the male gonads begin to differentiate under the influence of a cascade of testis-determining genes (reviewed in Refs. 42 and 43). In contrast, the recently characterized signaling molecule WNT-4 plays an active role in ovarian development (44). Ovaries are recognizable at about 10 weeks. If there is no secretion of anti-Müllerian hormone (AMH), a glycoprotein factor synthesized by the Sertoli cells of the testis (45), development of the Müllerian ducts proceeds and female internal structures—the Fallopian tubes, uterus, cervix and upper vagina—are formed (Figs. 3 and 4). In contrast, development of male genital structures derived from the Wolffian ducts, including the epididymis, ductus deferens, ejaculatory ducts, and seminiferous tubules, requires high local concentrations of testosterone secreted from

### Table 2. Characteristics of different forms of congenital adrenal hyperplasia

<table>
<thead>
<tr>
<th>Disease</th>
<th>21-Hydroxylase deficiency</th>
<th>11β-Hydroxylase deficiency</th>
<th>Aldosterone synthase deficiency</th>
<th>17α-Hydroxylase deficiency</th>
<th>3β-Hydroxysteroid dehydrogenase deficiency</th>
<th>Lipoid hyperplasia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defective gene</td>
<td>CYP21</td>
<td>CYP11B1</td>
<td>CYP11B2</td>
<td>CYP17</td>
<td>HSDB2</td>
<td>STAR</td>
</tr>
<tr>
<td>Alias</td>
<td>P450c21</td>
<td>P450c11</td>
<td>P450cald</td>
<td>P450c17</td>
<td>3β-HSD</td>
<td></td>
</tr>
<tr>
<td>Chromosomal location</td>
<td>6p21.3</td>
<td>8q24.3</td>
<td>8q24.3</td>
<td>10q24.3</td>
<td>1p13.1</td>
<td>8p11.2</td>
</tr>
<tr>
<td>Ambiguous genitalia</td>
<td>+ in ♂</td>
<td>+ in ♂</td>
<td>No</td>
<td>+ in ♂</td>
<td>+ in ♂</td>
<td>+ in ♂</td>
</tr>
<tr>
<td>Addisonian crisis</td>
<td>+</td>
<td>Rare</td>
<td>Salt wasting only</td>
<td>No</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Incidence (gen. pop.)</td>
<td>1:10–18,000</td>
<td>1:100,000</td>
<td>Rare</td>
<td>Rare</td>
<td>Rare</td>
<td>Rare</td>
</tr>
<tr>
<td>Hormones</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucocorticoids</td>
<td>↓</td>
<td>↓</td>
<td>Normal</td>
<td>Corticosterone normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mineralocorticoids</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Androgens</td>
<td>↑</td>
<td>↑</td>
<td>Normal</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Estrogens</td>
<td>Relatively ↓ in ♀</td>
<td>Relatively ↓ in ♀</td>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood pressure</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Na balance</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>K balance</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Acidosis</td>
<td>+</td>
<td>± Alkalosis</td>
<td>+</td>
<td>± Alkalosis</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Elevated metabolites</td>
<td>17-OHP</td>
<td>DOC, 11-deoxycortisol</td>
<td>Corticosterone, ±18-hydroxycorticosterone</td>
<td>DOC corticosterone, DHEA, 17Δ⁸Preg</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>(13)</td>
<td>(13)</td>
<td>(14)</td>
<td>(15)</td>
<td>(6)</td>
<td></td>
</tr>
</tbody>
</table>

17-OHP, 17-Hydroxyprogesterone; DOC, deoxycorticosterone; DHEA, dehydroepiandrosterone; 17Δ⁸Preg, 17-Δ⁸-hydroxypregnenolone.
but it involutes rapidly in the neonatal period (49). There-
weighing about the same as adult adrenals, or up to 10 g,
is approximately 10 times the size of the adult cortex,
cells that surround the fetal cortex. At term, the fetal cortex
formed in the 9th to 10th week by a second migration of
viability. The permanent, or adult, adrenal cortex is
DHEAS to produce estriol (48), a traditional marker of fetal
with an intact androgen receptor (Figs. 3 and 4)(46).
these questions remain unanswered.
B. Normal prenatal development of adrenal glands
The adrenal cortex is formed from mesoderm derived
from coelomic epithelium in the 4th week of gestation. By
6th to 7th week, steroids are secreted by the provisional
zonischloral zone, the functional cortex in fetal life (47). The provisional
cortex supplies DHEA sulfate to the fetal liver, where it
undergoes 16α-hydroxylation; the placenta utilizes 16α-
DHEAS to produce estriol (48), a traditional marker of fetal
viability. The permanent, or adult, adrenal cortex is
formed in the 9th to 10th week by a second migration of
cells that surround the fetal cortex. At term, the fetal cortex
is approximately 10 times the size of the adult cortex,
weighing about the same as adult adrenals, or up to 10 g,
but it involutes rapidly in the neonatal period (49). There-
C. Adrenarche
Beginning at 5–8 yr of age, there is an increase in the size
of the zona reticularis, correlating with a rise in serum
DHEAS and a modest increase in linear growth rate (56). This
process, termed adrenarche, occurs independently of
changes in ACTH, cortisol, or aldosterone production. Al-
though there has been speculation about a separate adrenal
androgen-stimulating hormone (57), no such factor has been
denoted mean values in normal infants.
vertical lines denote 95% confidence limits [adapted from Ref. 50]. Bottom, Timelines for
five aspects of sexual differentiation [adapted from Ref. 541]. Note
that 17-OHP levels are already markedly elevated in affected fetuses
during development of the external genitalia.

Leydig cells of the testis beginning at about 7 weeks; in the
absence of testosterone, Wolffian ducts regress. External
genital structures are also bipotential in early gestation and
differentiate as male under the influence of 5α-dihydro-
testosterone (reviewed in Ref. 17), which must interact
with an intact androgen receptor (Figs. 3 and 4)(46).

Anatomic hyperplasia of the adrenal is not seen invariably in
steroid 21-hydroxylase deficiency (54). The diagnostic utility of
ultrasound diagnosis of CAH may be improved, at least in
neonates, by examining not only size but also shape, surface
contours, and echogenicity (55). Steroid treatment can reverse
the structural abnormalities seen with sonography (55).

Anatomic and physiological data indicate that the hypotha-
lamic-pituitary-adrenal axis does not function until about the
eighth week of gestation. Experience with prenatal treatment of
CAH (see below, Section V.G), however, suggests that dexa-
methasone must be administered to the pregnant woman at risk
for an affected child as early as possible in the first trimester if
virilization of an affected girl is to be prevented. What, then, is
the mechanism for dexamethasone’s early action? Is there an-
other ACTH-independent glucocorticoid feedback pathway re-
sponsible for fetal adrenal steroid production? Could dexa-
methasone exert direct suppressive effects on the fetal adrenal?
These questions remain unanswered.

FIG. 2. Time course of prenatal sexual differentiation in male and
female fetuses. Top, Amniotic fluid levels of 17-OHP at various ages
of gestation. The scale is logarithmic. Open squares denote mean
values in fetuses affected with 21-hydroxylase deficiency, and closed
circles denote mean values in normal infants. Vertical lines denote
95% confidence limits [adapted from Ref. 50]. Bottom, Timelines for
five aspects of sexual differentiation [adapted from Ref. 541]. Note
that 17-OHP levels are already markedly elevated in affected fetuses
during development of the external genitalia.

The adrenal cortex is formed from mesoderm derived
from coelomic epithelium in the 4th week of gestation. By
the 6th to 7th week, steroids are secreted by the provisional
zone, the functional cortex in fetal life (47). The provisional
cortex supplies DHEA sulfate to the fetal liver, where it
undergoes 16α-hydroxylation; the placenta utilizes 16α-
DHEAS to produce estriol (48), a traditional marker of fetal
viability. The permanent, or adult, adrenal cortex is
formed in the 9th to 10th week by a second migration of
cells that surround the fetal cortex. At term, the fetal cortex
is approximately 10 times the size of the adult cortex,
weighing about the same as adult adrenals, or up to 10 g,
but it involutes rapidly in the neonatal period (49). There-
after, the permanent cortex assumes the steroidogenic
functions and develops the three-zoned organization of
the adult gland. Several transcription factors are known to
be critical for adrenal development. Steroidogenic factor-1
(SF-1, also called Ad4BP, reviewed in Ref. 50), induces
genes involved in steroid synthesis, and is in turn nega-
tively regulated by DAX-1, the gene affected in congenital
adrenal hypoplasia (51). Human fetal adrenal development is
regulated primarily by fetal pituitary ACTH. ACTH is
not a mitogen per se; rather its actions on the fetal adrenal
cortex are mediated in autocrine/paracrine fashion by
several growth factors. In cultured human fetal adrenal
cortical cells, epidermal growth factor (EGF), basic fibro-
blast growth factor (bFGF), human CG (hCG), and insulin-
like growth factors I and II (IGF-I and -II) are mitogenic,
whereas activin and transforming growth factor-β (TGFβ)
inhibit proliferation. IGF-II, activin, and TGFβ also mod-
ulate ACTH-stimulated steroidogenesis (reviewed in Ref.
49). In the absence of ACTH, as in anencephaly, the fetal
adrenal involutes in the second trimester (52).

In addition, catecholamines and neuropeptides secreted
by the adrenal medulla, as well as direct innervation of the
adrenal cortex, may influence development of the cortex
(reviewed in Ref. 53).
D. Prenatal virilization

Adrenal secretion of excess androgen precursors does not significantly affect male sexual differentiation. In females affected with CAH, however, the urogenital sinus is in the process of septation when the fetal adrenal begins to produce excess androgens; levels of circulating adrenal androgens are apparently sufficiently high to prevent formation of separate vaginal and urethral canals. Further interference with normal female genital anatomy occurs as adrenal-derived androgens interact with genital skin androgen receptors and induce clitoral enlargement, promote fusion of the labial folds, and cause rostral migration of the urethral/vaginal perineal orifice. However, internal Wolffian structures, such as the prostate gland and spermatic ducts, are usually not virilized, presumably because development of the Wolffian ducts requires markedly higher focal concentrations of testosterone than the external genitalia. This is supported by animal studies showing that unilateral castration causes ipsilateral mesonephric duct involution (61). Nevertheless, severely affected females may occasionally have some development of typically male internal genital structures; carcinoma of prostate tissue has been reported in an affected female (62).

Thus, the typical result in severely affected girls is ambiguous or male-appearing external genitalia with perineal hypospadias, chordee, and undescended testes (63). The severity of virilization is often quantitated using a five-point scale developed by Prader (Figs. 3 and 4) (64). Not all classic CAH females develop the same degree of genital ambiguity. One might speculate that the physical signs of androgen
excess are dependent not only on direct adrenal secretion of androgen precursors, but also on the efficiency with which such hormones are converted to more potent products, such as dihydrotestosterone by peripheral enzymes such as 5α-reductase (17). Additionally, the concentration (65) and transcriptional activity (66) of androgen receptors (both of which are influenced by a highly polymorphic CAG repeat sequence within the coding region) may play a further role in determining genital phenotype.

As another presumed effect of prenatal exposure to excess androgens, both male and female affected infants are longer than average at birth (67). Moreover, infant girls with CAH have higher than typical LH levels—into the range expected for healthy infant boys—presumably due to exposure to higher than normal levels of prenatal androgens and/or other sex hormones (68). Although there have been no studies of LH pulsatility or responsiveness to GnRH in infants with CAH, women with well controlled classic CAH, but not those with nonclassic CAH, have exaggerated LH responses to GnRH and increased production of ovarian androgens. This is consistent with the idea that early exposure to androgens or progestins causes permanent abnormalities in the hypothalamic-pituitary-gonadal axis in CAH women (69) (see Section III.G).

E. Salt wasting

Among classic CAH patients, about three-fourths cannot synthesize adequate amounts of aldosterone due to severely impaired 21-hydroxylation of progesterone. Aldosterone is essential for normal sodium homeostasis; deficiency of this hormone results in sodium loss via the kidney, colon, and sweat glands (70).

Severely affected patients invariably have concomitant cortisol deficiency that exacerbates the effects of aldosterone deficiency. Glucocorticoids normally increase cardiac contractility, cardiac output, sensitivity of both the heart and the vasculature to the pressor effects of catecholamines and other pressor hormones, and work capacity of skeletal muscles (71). In the absence of glucocorticoids, cardiac output decreases. This decreases glomerular filtration, leading to an inability to excrete free water and consequent hyponatremia. Thus, shock and severe hyponatremia are much more likely in 21-hydroxylase deficiency, in which both cortisol and aldosterone biosynthesis are affected, than in (for example) aldosterone synthase deficiency, in which only one biosynthetic pathway is impaired (72).

Although catecholamine secretion has not, to our knowledge, been studied in patients with CAH, high levels of glucocorticoids are required for normal development of the adrenal medulla and for expression of the enzymes required to synthesize catecholamines (73). Indeed, mice with 21-hydroxylase deficiency exhibit abnormal development of the adrenal medulla and secrete reduced levels of catecholamines (74). Catecholamine deficiency could further exacerbate the shock engendered by glucocorticoid and mineralocorticoid deficiency.

In addition, accumulated steroid precursors may directly antagonize the mineralocorticoid receptor and exacerbate mineralocorticoid deficiency, particularly in untreated patients (75). Progesterone is well known to have antimineralocorticoid effects (76–79), and it and/or a metabolite are likely culprits in this phenomenon. However, there is as yet no evidence that 17-OHP has direct or indirect antimineralocorticoid effects.

Salt wasting may include such nonspecific symptoms as poor appetite, vomiting, lethargy, and failure to gain weight. Severely affected patients with CAH usually present at 1–4 weeks of age with hyponatremia, hyperkalemia, hyperreninemia (see Section IV.C.2) and hypovolemic shock. These “adrenal crises” may prove fatal if proper medical care is not delivered. This problem is particularly critical in infant boys who have no genital ambiguity to alert physicians to the diagnosis of CAH before the onset of dehydration and shock (80). The mortality rate for CAH remains high in such patients, as suggested by the relative paucity of male patients identified in case reports (81). It is for this reason that many states in the United States and a number of countries have adopted newborn screening for CAH (see Section IV.B).

The rapidity of onset and severity of a salt wasting crisis may reflect the individual’s ancillary homeostatic mechanisms for sodium and fluid conservation. Such factors might include the concentration and transcriptional activity of mineralocorticoid receptors in the kidney and elsewhere, and the ability to increase vasopressin or decrease atrial natriuretic factor (82) in response to volume contraction.

Siblings may be discordant for salt wasting (83). Furthermore, CAH patients known to have severe salt wasting episodes in infancy and early childhood may show improved sodium balance and relatively more efficient aldosterone synthesis with age. Unrelated individuals carrying identical mutations may manifest different degrees of salt wasting (84). Although explanations for these observations are not immediately apparent, both genetic and nongenetic factors may contribute to the presence or absence of the salt wasting trait. Extraadrenal 21-hydroxylase has been detected by in vivo metabolic studies (85), but molecular genetic investigation has yielded contradictory results as to whether CYP21 could be a source for this activity (86–89). Other enzymes with 21-hydroxylase activity have not been identified in humans, although such enzymes have been identified in rabbit liver (90).

F. Postnatal signs of androgen excess

Ongoing adrenal sex steroid production in the untreated or incompletely treated patient causes several problems. Boys have inappropriately rapid somatic growth with advancement of epiphyseal maturation, although this may not be apparent in the first 18 months of life (91). Pubic hair and apocrine body odor develop, and penile size increases without testicular enlargement.

Girls may show similar signs of sex steroid excess as well as progressive clitoral enlargement. In adolescence, poorly controlled girls manifest acne, hirsutism, and ovarian dysfunction (see below).

There is considerable interindividual variation in pre- and postnatal signs of androgen excess. This may be attributed directly to differences in the absolute levels of androgen precursors secreted by the affected adrenals, or to the efficiency of conversion of precursors to more potent androgens.
Alternatively, variations in androgen receptor expression or activity may contribute to phenotype. For example, expansion of the CAG repeat sequence in exon 1 results in decreased androgen receptor transactivation of target DNA sequences (66). In a correlative study, higher hirsutism scores correlated with fewer CAG repeats in women with idiopathic hirsutism (92).

Although childhood somatic growth is excessive in CAH patients (67), adult height is often suboptimal compared with the surrounding healthy population and with parentally determined target height (93–99). Whereas untreated patients grow rapidly, patients treated with excessive doses of glucocorticoids may suffer growth retardation. This is discussed below in Section V.A.

Although androstenedione is elevated, DHEAS is suppressed in CAH children (59). This is likely due to exogenous glucocorticoid suppression of the adrenal (60). An adrenal androgen-stimulating hormone (AASH) separate from ACTH has been postulated (100), but this hypothesis cannot be tested in any definitive manner based on data from CAH patients.

G. Reproductive function in classic CAH

1. Females. Reproductive problems for women with CAH become apparent in adolescence. The average age at which menarche occurs in inadequately treated girls is late compared with healthy peers (101). Such girls and women with CAH often have a clinical picture similar to polycystic ovarian syndrome with sonographic evidence of multiple cysts, anovulation, irregular bleeding, and hyperandrogenic symptoms (102). Moreover, a significant reduction in insulin sensitivity, although not clinical diabetes, is found among young women with nonclassic CAH as compared with controls of similar age and weight (103).

The basis for these problems is not precisely known. Several hypotheses have been advanced: 1) Hypothalamic aromatization of excess adrenal androstenedione might interfere with LH-releasing hormone secretion (104). 2) Excess adrenal progesterone might act as a “mini-pill” to inhibit normal cyclicity (101), or it might antagonize estrogen effects (101, 105). 3) Elevated progestins or sex steroids could induce abnormal ovarian function by programming the hypothalamus early in development (69). 4) Androgen excess might directly damage the ovaries. 5) Adrenal rest tissue might displace normal gonadal parenchyma.

The majority of women with CAH eventually undergo menarche. In general, the regularity of menses depends on the adequacy of treatment. A small proportion of women do not undergo menarche and are unable to suppress progestin levels even when 17-OHP is adequately suppressed (105).

Furthermore, breast development is suppressed in females with CAH. Evidence from animal studies suggests that testosterone exposure in utero may also suppress the breast anlage, resulting in poor breast development at adolescence (106). However, this problem is apparently due mainly to the combined effects of androgen excess and cortisol deficiency, because it is reversible with treatment (107, 108).

Pregnancy outcome in classic CAH has been recently reviewed. During pregnancy, women are optimally managed with hydrocortisone or prednisone (109, 110). Due to pregnancy-induced alterations in steroid metabolism and clearance, doses need be increased compared with doses used in nonpregnant women with CAH. It should be recognized that in this situation, treatment is directed at the mother and not at the fetus, for hydrocortisone and prednisone do not effectively cross the placenta. Interestingly, despite elevated maternal testosterone of 400–600 ng/dl, unaffected female offspring appear to have no genital virilization (109). Apparently, placental aromatase effectively prevents maternal androgens from reaching the fetus. Elevated maternal sex hormone-binding globulin (111) and androgen antagonism by progesterone (112) may also restrict transplacental passage of androgens.

There is no evidence of an excess of congenital malformations in offspring of women with CAH.

2. Males. Men with CAH less frequently have impaired gonadal function compared with affected women. Most affected males are able to father children or at least have normal sperm counts (113). Low sperm counts, when they occur, do not always preclude fertility (114). Among simple virilizing patients, testicular integrity may be normal even in the absence of treatment (115). A prominent complication in CAH males is the development of testicular adrenal rests (116). This is discussed in Section III.I.2.

H. Neuropsychology of CAH

1. Cognitive effects. Although there have been occasional reports of elevated IQ among CAH patients (117), this has not been generally observed. To the contrary, salt wasting patients who suffer hyponatremic dehydration and shock may sustain permanent brain injury with resultant lower cognitive test scores (118–120). Certain sexually dimorphic cognitive abilities, such as spatial abilities, may be enhanced among CAH girls (121–123). Females with CAH are more likely to be left-handed (as are males) (124) but do not differ from unaffected women in degree of cerebral lateralization (125). Magnetic resonance imaging showed white matter abnormalities in the brains of CAH patients more often (117, 126) than in controls in two of three recent studies (127). Thus, neurodevelopmental evaluation is warranted in children with CAH. Patients who have experienced severe hyponatremia should be considered for enrollment in early intervention programs if neurodevelopmental milestones are delayed.

2. Effects on gender role and identity. The influence of prenatal sex steroid exposure on personality is controversial (reviewed in Refs. 128–132); also see Sections V.D and V.E.). In considering this question, it is important to distinguish between gender role, sexual orientation, and gender identity. Gender role refers to gender-stereotyped behaviors such as choice of play toys by young children. Parents of young girls with CAH often report that their daughters prefer to play with trucks as compared with dolls. Indeed, low interest in maternal behavior, beginning with infrequent doll play in early childhood and extending to lack of interest in child.
rearing for older girls and women, is a recurring theme in CAH research (133–136). Some investigators have noted tomboyish (137) or aggressive (138) behavior among girls with CAH or a male pattern of distance in social relations (139). Others have found that young patients do not differ significantly from controls for nine parameters of psychopathology including aggression and hyperactivity (140); older girls or women with CAH have tested similarly (133, 141). The amygdala, an androgen-sensitive brain center controlling fear and aggression, is smaller by MRI among children with CAH (142); these MRI-based structural differences have not yet been directly correlated with psychological testing.

Sexual orientation refers to homosexual vs. heterosexual preferences. In most studies, a small but significant percentage of adult women with CAH have been actively homosexual or bisexual or have an increased tendency to homoerotic fantasies (143–145). These characteristics occur more frequently in women with the salt wasting form of 21-hydroxylase deficiency. A review of German patients found no increase in homosexuality among affected women but did find a decreased frequency of marriage and childbearing, suggesting more general psychosocial dysfunction among patients (146).

Gender identity refers to self-identification as male or female. Spontaneous gender reassignment back to male has been reported in cases of males with penile trauma who were raised as females (147, 148) or male pseudohermaphrodites raised as females, especially in cases of 5α-reductase or 17α-hydroxylase deficiencies, in which the brain may be exposed to high circulating levels of androgens (reviewed in Ref. 132). Conversely, female-to-male transsexuals may have relatively high levels of androgens and a high incidence of polycystic ovary syndrome (149). However, self-reassignment to the male sex is unusual in women with CAH (145, 150). When it occurs, it may be related to delays in gender assignment or genital surgery or to inadequate suppression of adrenal androgens with glucocorticoid therapy (151). Severely virilized females are more likely to be raised as males in cultures that value boys more highly and/or in third world countries in which the diagnosis is likely to be delayed (152–154). There have been few studies directly comparing psychosexual functioning in severely virilized genetic females with CAH raised as women or men, but it does not appear that those raised as men are psychologically better adjusted than those raised as women (155).

The uncertainty concerning the effects of prenatal and postnatal effects of androgen on gender identity and gender role extends not only to the female CAH population, but also to male pseudohermaphrodites of other etiologies. The role of external genital anatomy before and after genital surgery in fueling problems relating to gender is unclear compared with the roles of prenatal hormone exposure, rearing by the family, and community attitudes. Unfortunately, much of the data in this area are anecdotal (reviewed in Refs. 130 and 131). These issues are discussed further in Section V.D.

In summary, most CAH children manifest normal neuro-psychological development. Moreover, despite a tendency toward male gender role behavior and homoerotic fantasy, most girls with CAH identify as females and exhibit heterosexual preference.

I. Tumors

1. Adrenal. Almost 60% of patients with incidentally discovered adrenal masses (incidentalomas) have exaggerated 17-OHP responses to ACTH stimulation; the frequency of abnormal responses is even higher in patients with bilateral adrenal masses (156). The frequency of germline mutations in CYP21 in such patients is low (157). However, the incidence of adrenal masses appears to be higher in CAH patients and in heterozygotes than in the general population (158). Histological types of adrenal tumor include adenoma, myelolipoma (159, 160), and hemangioma (161). Steroid-responsive hyperplastic adrenal nodules can present in previously undiagnosed patients late in life and can potentially be confused with virilizing adrenal adenomas (159, 162, 163). Because these tumors may regress with glucocorticoid therapy, it may be unnecessary to resect them if they are carefully followed. Rarely, virilizing adrenal carcinoma has been found in CAH patients (164, 165), but most adrenal masses in children with CAH are benign (166).

Partially autonomous cortisol secretion is rare in adrenal adenomas arising in patients with hormonal evidence of 21-hydroxylase deficiency (167). Acute adrenal insufficiency may develop after resection of such a nodule if steroids secreted by the nodule have suppressed ACTH secretion, leading to atrophy of the remaining adrenal cortex (168).

2. Testicular. Although seen most often in inadequately treated patients, testicular adrenal rests accompanied by deficient spermatogenesis may occur despite treatment, particularly in males with the salt wasting form of 21-hydroxylase deficiency (104, 169, 170). These tumors, although most often benign, have prompted biopsies and sometimes even orchietomy (171). The preferred mode of treatment consists of effective adrenal suppression with dexamethasone, since many of these tumors are ACTH responsive. When they do not respond to dexamethasone, testis-sparing surgery may be performed after imaging of the tumor by sonography and/or MRI (171). Adrenalectomy (Section V.C.2) would not be expected to alleviate problems caused by gonadal adrenal rests (172). Testicular masses have been detected in children as young as 3 yr with CAH (116, 173), prompting the recommendation that boys undergo a baseline testicular sonogram by adolescence (174). The testes of affected males should be carefully examined throughout childhood, adolescence, and adulthood.

The main differential in the diagnosis of a virilizing testicular mass is a Leydig cell tumor. Such tumors can occasionally secrete high levels of 17-OHP suggesting CAH, but the secretion of 17-OHP from such a tumor will not be suppressed by dexamethasone or stimulated by ACTH (175). In general, bilateral tumors or those that decrease in size with dexamethasone are very likely to be adrenal rests. An adrenal rest may also be diagnosed if selective spertmatic vein catheterization to assay steroids produced by the testis reveals high levels of 11β-hydroxylated steroids (e.g., 21-deoxycortisol) (176), because the 11β-hydroxylase enzyme is not active in testicular tissue.

3. Pituitary. Although glucocorticoid replacement doses exceed physiological cortisol production, CRH and ACTH of-
ten are not fully suppressed by treatment as evidenced by basal levels and by stimulation testing with ovine CRH. In one study, four of seven CAH patients undergoing MRI of the head showed pituitary abnormalities (three apparent microadenomata and one empty sella) (177). However, to the best of our knowledge symptomatic pituitary tumors have not been reported.

J. Nonclassic CAH phenotypes

1. Signs. Patients with the mild, nonclassic form of 21-hydroxylase deficiency may have any of the signs of postnatal androgen excess listed above, but affected females are born with nonambiguous (normal or with mild clitoromegaly) external genitalia. Adrenal steroid precursors of 21-hydroxylase are only mildly elevated in nonclassic CAH and are intermediate between those of heterozygote carriers of the enzyme deficiency and those who are severely affected (35). Depending on the laboratory, affected individuals have serum 17-OHP levels of greater than 1,000 or 1,500 ng/dl 60 min after an intravenous bolus of cosyntropin (ACTH 1–24). Due to circadian variability of adrenal cortical hormones (178), the diagnosis may be missed by measuring only baseline serum 17-OHP late in the day. The severity of signs and symptoms of mild androgen excess varies widely, and probably many affected individuals are asymptomatic. The most common presenting symptoms are premature pubarche in children (179, 180), or severe cystic acne (181), hirsutism, and oligomenorrhea in young women (58, 182).

Nonclassic male patients diagnosed after puberty have presented with acne or infertility, but are most often diagnosed in the course of family studies and are entirely asymptomatic (183, 184). Rarely, a nonclassic male has presented with unilateral testicular enlargement (183). Precise clinical distinction between classic simple virilizing disease and the nonclassic disorder is often difficult among boys, since 17-OHP levels form a continuum between the mild and severe cases, and signs of androgen excess are much less apparent than in females.

Aldosterone synthesis and sodium balance are not compromised to any clinically significant extent in patients with nonclassic 21-hydroxylase deficiency (185), although under stress conditions subtle abnormalities may be elicited (186). Likewise, cortisol synthesis during stress is not impaired to any clinically significant degree (185), and there have been no deaths from adrenal insufficiency reported with this condition.

There are conflicting reports as to whether adult height is compromised in nonclassic CAH. Height sd scores were lower in one study compared with the population (−0.99 ± 0.98) but not when compared with midparent heights (0.43 ± 0.77) (94). Similarly, other investigators found no differences between nonclassic patients and their unaffected siblings (187).

The pathophysiology of the less frequent and milder reproductive problems associated with nonclassic CAH is presumably similar to that suggested for classic CAH. Data regarding reproductive function in nonclassic CAH come mainly from studies of populations referred for symptoms and signs of hyperandrogenism and/or infertility; ascertainment bias obviously affects such studies. In one report 39% of women presented with hirsutism, 39% with oligomenorrhea or other signs of polycystic ovaries, and 22% with no obvious signs of androgen excess (188). Based on data derived from family studies, it is clear that many individuals with mild 21-hydroxylase deficiency have minimal symptoms and are not brought to medical attention. French investigators found that half of the patients in their clinic became pregnant before the diagnosis of nonclassic CAH was made. All the others who desired pregnancy successfully conceived during hydrocortisone treatment; only 1 of 20 women required clomiphene citrate to conceive (189). Clomiphene without hydrocortisone has also successfully induced ovulation (190). For women who have conceived without hydrocortisone treatment, it is not necessary to initiate therapy during pregnancy; testosterone levels in nonpregnant women with nonclassic CAH are generally lower than typical testosterone levels in normal women during the second trimester of pregnancy (i.e., less than about 150 ng/dl).

2. Incidence. Because the signs of androgen excess in nonclassic 21-hydroxylase deficiency can be difficult to discern, particularly in males, the most reliable estimates of allele and disease frequencies come from ascertainment of affected individuals in the course of studies of kindreds in which classic and nonclassic 21-hydroxylase deficiency are segregating (191–193). The disease frequency is estimated at 0.1% of the general population but it occurs in 1–2% of Hispanics and Yugoslavs and 3–4% of Ashkenazi (Eastern European) Jews. Similar frequencies have been estimated from a small screening study using morning salivary 17-OHP levels (194). In New Zealand, molecular screening of normal newborns showed that 5% are carriers for mutations in the 21-hydroxylase gene (CYP21) associated with either classic or nonclassic 21-hydroxylase deficiency (see Section VII.F). This implies a disease frequency for nonclassic 21-hydroxylase deficiency of 0.06%, in good agreement with estimates in the general American population (195).

Although it has been suggested that nonclassic 21-hydroxylase deficiency represents the most frequent autosomal recessive genetic disorder in man (191, 194), the proportion of affected individuals who have problems with androgen excess is not known. To the best of our knowledge there has been no prospective study of symptomatology in any nonclassic 21-hydroxylase deficiency patient population. Because of the stigma and anxiety that may be associated with the diagnosis of a genetic disease, we suggest that nonclassic 21-hydroxylase deficiency be initially considered a "genetic polymorphism" and discussed as a disease only if signs of androgen excess develop.

Conversely, only a small percentage of individuals presenting with signs of androgen excess prove to be affected with nonclassic 21-hydroxylase deficiency. Among children referred for precocious pubarche, 4–7% have nonclassic 21-hydroxylase deficiency (180, 196–198). Among 31 women referred for acne, none had 21-hydroxylase deficiency, although a majority of patients had exaggerated adrenal responsiveness to ACTH stimulation (199). In
the largest study of hyperandrogenic women, only 6% of 400 hirsute French women had hormonal profiles compatible with the diagnosis of late-onset CAH (200). These statistics have been borne out in other large clinic population samples (201–203). The lowest incidence of nonclassic CAH was 1.2% in 83 hyperandrogenic Californian women (204), whereas the highest incidence of nearly 14% was detected in New York women (205). Variations in the frequency of nonclassic alleles among different ethnic groups may account for some of the discrepancies noted. New York has a high proportion of Ashkenazi Jews, and this group has the highest frequency of the typical nonclassic CYP21 allele, valine-to-leucine 281 (V281L), associated with HLA-B14,DR1 (see Section VI.E) (191, 206).

There does not appear to be a high prevalence of nonclassic CAH in men with infertility (207), and, conversely, most men with nonclassic CAH ascertained through family studies have proved fertile. However, oligospermia and infertility have occasionally been described (104, 183, 208). In some cases, these problems may be reversed by glucocorticoid treatment (104, 208).

The incidence of classic CAH is discussed in Section IV.B.

K. Heterozygotes

Heterozygotes carrying a single mutant allele have slightly elevated 17-OHP levels after ACTH stimulation, but there is substantial overlap with unaffected individuals (35). The range of most heterozygotes’ 17-OHP response at 60 min after cosyntriopin stimulation is approximately 200–1,000 ng/dl (209). Of 53 women with signs of hyperandrogenism who were suspected by hormonal testing of being carriers for CYP21 mutations, such mutations could be detected in only 37; in contrast, mutations could be detected on both alleles in 15/16 women who had 17-OHP or 21-deoxycortisol levels in the range expected for nonclassic CAH (210).

In view of these problems with hormonal detection of heterozygotes, genotyping would appear to be a superior heterozygote detection method. This is discussed further in Sections IV.B and VI.H.

Mothers of children with classic CAH are no more likely to show signs of androgen excess than age, sex, and BMI-matched controls (211). However, children referred to an endocrine clinic for premature pubarche or hirsutism showed a higher prevalence of heterozygous CYP21 mutations compared with 80 adult controls who were not screened for hyperandrogenic signs or symptoms (212). Since potential heterozygotes in the latter study were culled from a symptomatic referral population, they may not represent the population-at-large carrying CYP21 mutations. With estimated nonclassic and classic heterozygote frequencies of 10% (191) and 1.5% in the general population, respectively, it is unlikely that heterozygosity confers a clinically significant reproductive disadvantage. Screening of men referred for evaluation of infertility has not revealed a high prevalence of nonclassic 21-hydroxylase deficiency patients or heterozygotes (207).

![Fig. 5. Simplified flowchart for initial evaluation of ambiguous genitalia.](https://academic.oup.com/edrv/article-abstract/21/3/245/2423831/12 March 2019)
as 5α-reductase deficiency. A team consisting of neonatologist, pediatric endocrinologist, urologist, and preferably an experienced social worker and/or child psychiatrist should promptly review the essential early diagnostic data and make a recommendation to the family as to the sex of rearing and any medical and/or surgical treatments. These recommendations should be based on both the current state of knowledge of psychosexual development in intersex individuals (discussed in Section III.H) and the feasibility of surgical correction (Section V.D). Although all available options should be reviewed with the family, these recommendations should be as unequivocal as possible.

B. Newborn screening

CAH is a disease well suited to newborn screening since it is a common and potentially fatal childhood disease, it is easily diagnosed by a simple hormonal measurement in blood, and early recognition and treatment can, in principle, prevent serious morbidity and mortality (Fig. 6).

1. Technical considerations. The diagnosis of CAH is suspected when one finds a markedly elevated filter paper blood 17-OHP level by RIA (215, 216); normative values for filter paper assays vary in different laboratories. These assays use the same “Guthrie” cards as are used for screening for phenylketonuria and hypothyroidism. Subsequent measurement of serum 17-OHP is usually performed to confirm the diagnosis.

Premature, sick, or stressed infants tend to have higher levels of 17-OHP than term infants and generate many false positives unless higher normal cut-offs are used (Fig. 7). Suggested weight-adjusted cut-offs range from 165 ng/ml for infants under 1,300 g to 40 ng/ml for infants over 2,200 g in Wisconsin (217); in Texas, cut-offs of 40 and 65 ng/ml are used for infants greater or less than 2,500 g, respectively (218). Elevated 17-OHP levels in preterm infants have been confirmed by HPLC and are thus not due to cross-reaction with other steroids [however, some 17-OHP RIAs do cross-react with other steroids; these include 15β-hydroxylated compounds, which are apparently generated by gut bacteria and resorbed through the enterohepatic circulation (219)]. The steroid profiles in preterm infants suggest a functional deficiency of several adrenal steroidogenic enzymes with a nadir in function at 29 weeks of gestation (220).

2. Incidence and cost effectiveness. As determined by screening (Table 3, and summarized in Ref. 81) the highest incidence of classic CAH occurs in two geographically isolated populations, the Yupik Eskimos of Western Alaska (1:280) (221) and the French island of La Reunion in the Indian Ocean (1:2,100). The incidence in most other populations ranges from approximately 1:10,000 to 1:18,000 (81, 195, 217, 218, 222–224).

It is now well established that screening markedly reduces the time to diagnosis of infants with CAH (218, 224–226). The main putative benefit of this is reduced morbidity and mortality because infants with salt wasting disease are diagnosed more promptly. As undiagnosed infants who die suddenly may not be ascertained, it is difficult to demonstrate a benefit of screening by direct comparison of death rates from CAH in unscreened and screened populations. However, males...

---

![Flowchart for decisions pertaining to newborn screening for 21-hydroxylase deficiency. ACTH stim 17-OHP, 17-Hydroxyprogesterone level after cosyntropin stimulation; 'lytes, electrolytes.](image)

![Levels of 17-OHP in dried blood samples from the Wisconsin neonatal screening program, plotted against birth weight. The heavy line represents mean values and the dotted lines represent 95% confidence limits. The heavy dashed line denotes threshold notification values in the Wisconsin program for infants of various birth weights. Adapted from Ref. 217.](image)

---

**TABLE 3. Frequency of classic 21-hydroxylase deficiency determined from neonatal screening (representative populations)**

| Region                      | Incidence | No. detected/ no. screened | Reference*
|-----------------------------|-----------|-----------------------------|-----------
| Alaska, Yupik Eskimos       | 1:280     | 5/1,131                     | (221)     |
| France, La Reunion          | 1:2,100   | 7/14,987                    | (81)      |
| Sweden                      | 1:9,800   | 73/not given                | (225)     |
| United States, Wisconsin    | 1:11,000  | 14/149,684                  | (217)     |
| France, Lille               | 1:13,000  | 31408138                    | (539)     |
| Japan                       | 1:18,000  | Not given/4,500,000         | (232)     |
| United States, Texas        | 1:16,000  | 1211,936,998                | (218)     |
| Scotland                    | 1:17,000  | 7/119,960                   | (81)      |
| Italy                       | 1:18,000  | 27/420,960                  | (224)     |
| New Zealand                 | 1:23,000  | 23/536,915                  | (223)     |

*a References to earlier studies are found in Ref. 81.*
with salt wasting CAH are more likely than females to suffer from delayed or incorrect diagnosis because there is no genital ambiguity to alert the clinician. Thus, a relative paucity of salt wasting males in a patient population may be taken as indirect evidence of unreported deaths from salt wasting crises. Indeed, females outnumbered males in some (227, 228) but not all (229) retrospective studies in which CAH was diagnosed clinically. In contrast, cases of salt wasting CAH ascertained through screening programs are equally or more likely to be males rather than females (224–226).

As regards morbidity, infants ascertained through screening have less severe hyponatremia (225) and tend to be hospitalized for shorter periods of time (although the difference falls short of statistical significance) (226).

Although salt wasting males would seem to derive the greatest benefit from screening programs, the delay before correct sex assignment of severely virilized females is also markedly reduced (79, 81, 225). Moreover, males with simple virilizing disease may otherwise not be diagnosed until rapid growth and accelerated skeletal maturation are detected later in childhood, at which time final height may already be adversely affected. However, it is debatable whether this last benefit itself justifies the costs of a screening program.

The estimated cost of screening each newborn infant was $2.70 in Sweden. With a disease incidence of 1:9,800 in this population, 102 affected newborns are expected per million infants screened; 51 should be males, of whom 75% should be salt wasters. The total cost for screening each million infants is $2,700,000, and thus the cost for each of 38 salt wasting males expected to be detected by screening is $71,000. The cost of newborn CAH screening in Texas was higher at $87,000 per CAH case, as separate hormonal assays were performed on each infant at birth and again at 1–2 weeks of age (218). All infants with salt-wasting CAH were detected on the first screen, so that the second screen may not be cost effective (230).

Nevertheless, these costs are within the general range estimated for other newborn disease detection programs. By comparison, targeted newborn screening for hemoglobinopathy in Alaska cost approximately $200,000 per death averted (231).

Patients with nonclassic 21-hydroxylase deficiency are occasionally detected by newborn screening. In Texas, 87% are detected on the second of the two routine screening tests, with an overall frequency of nonclassic disease of 1:35,870 (218). This is much less than the 1:1,000 frequency in the general population estimated from nonclassic allele frequencies in kindreds in which classic 21-hydroxylase deficiency is segregating (191, 192). Thus, neonatal screening using hormonal assays is not an efficient way to detect nonclassic disease. As yet, there are no follow-up studies of patients with nonclassic disease who have been ascertained by neonatal screening to determine how frequently they develop signs of androgen excess. There have also not been any systematic genotyping studies of nonclassic patients ascertained through neonatal screening to determine whether their genotypes differ from nonclassic patients ascertained through family studies or because they had developed signs of androgen excess. In a small study in Japan, all four patients ascertained through neonatal screening were compound heterozygotes for classic mutations (232), consistent with the higher 17-OHP levels seen after cosyntropin stimulation in compound heterozygotes (36)(also see Section VI.I). These data suggest that infants who are homozygous for mild CYP21 mutations are less likely to be detected by basal hormone screening.

3. Strategies for follow-up. To obtain adequate sensitivity, the cut-off levels for 17-OHP are typically set low enough that 0.3–0.5% of all tests are reported as positive. Therefore, specificity is only 2%, i.e., 98% of all positive tests are false. The above estimates for cost of detection do not include costs for follow-up of false positives. In Texas (218), both the infant’s primary physician and a pediatric endocrinologist are notified of all positive screens. Moderately elevated 17-OHP levels (40–100 ng/ml for term infants) are followed up with a repeat filter paper specimen. Higher values are evaluated with electrolytes and a serum 17-OHP level; if these are not unequivocally normal, the infant is then referred to a pediatric endocrinologist. A cosyntropin stimulation test is then usually performed.

4. Molecular genetic screening. Much of the expense of following up positive newborn screening tests could be avoided with a second level of screening based on detection of actual mutations (see Section VI.H). This could be accomplished on DNA extracted from the same dried blood spots as are used for hormonal screening. Because 90–95% of mutant alleles carry one or more of a discrete number of mutations (see Section VI.F), samples that carry none of these mutations may be presumed with more than 99% confidence to be unaffected. Heterozygous carriers of a mutation for classic 21-hydroxylase deficiency would still need to be followed up due to the chance that the other allele might carry a novel mutation, but less than 2% of individuals are carriers of classical 21-hydroxylase deficiency alleles.

Two large-scale studies of the utility of genotyping in screening programs have shown that this is a useful adjunct to hormonal measurements (195, 233). One study examined cost and found it to be approximately $5 per sample analyzed (195). At present, however, there are few laboratories equipped to do rapid, accurate, and large-scale CYP21 genotyping.

C. Further biochemical evaluation

1. The cosyntropin stimulation test. As previously mentioned, a basal serum or filter paper 17-OHP may not be fully informative, and it may be necessary to evaluate the patient further. In cases where there is no newborn screening program, but one suspects CAH based on ambiguous genitalia, cosyntropin stimulation should be deferred beyond the first 24 h of life. There is a high incidence of both false-positive and false-negative results when samples are obtained immediately after birth. Another justification for performing stimulation testing is that 17-OHP may be elevated in other enzymatic defects, e.g., 11β-hydroxylase or 3β-hydroxysteroid dehydrogenase deficiencies. Ideally, to fully differentiate the various enzymatic defects, the clinician should measure 17-OHP, cortisol, DOC, 11-deoxycortisol, 17-OH-pregnenolone, DHEA, and androstenedione at 0 min and 60 min (Fig. 8). If blood volume is an issue in small
infants, a sample is collected only at 60 min. Precursor to produce ratios are particularly useful in distinguishing the different enzymatic defects. If the diagnosis remains unclear, it may be desirable to test the child and later rest after partially or completely tapering glucocorticoids. Our practice is to use a uniform dose of 0.25 mg cosyntropin, providing a pharmacological stimulus to the adrenal cortex. This diagnostic test should be distinguished from the low-dose ACTH stimulation test now becoming increasingly popular for evaluating the integrity of the hypothalamic-pituitary-adrenal axis (234, 235).

2. Evaluation of salt wasting. Elevated PRA values, and particularly the ratio of PRA to 24 h urinary aldosterone, are often used as markers of impaired aldosterone synthesis (236). They can also be increased in patients with normal aldosterone secretion who have high circulating levels of ACTH, 17-OHP, and progesterone, making poorly controlled simple virilizers biochemically resemble salt wasters. Conversely, mineralocorticoid therapy may aid adrenal suppression in such patients (236). Ideally, plasma and urinary aldosterone levels should be correlated with PRA and with sodium balance to gain an accurate assessment of phenotype. A direct immunoradiometric assay of active renin may be an alternative to PRA measurements, with the advantage of smaller sample requirements, but it is not yet widely available (237). In interpreting renin levels, it should be kept in mind that they are normally higher in neonates than in older children, and age-specific reference values for both immunoactive renin (238) and for PRA (239) in infants and children vary by laboratory.

3. Other hormones useful in diagnosis and monitoring of CAH. Several other diagnostic biochemical assays have been proposed, but few are widely available. Assays of 21-deoxycortisol can detect more than 90% of CAH heterozygotes (39, 40). Levels of an androgen metabolite, 3α-androstanediol glucuronide, are elevated in nonclassic 21-hydroxylase deficiency (240) and highly correlated with levels of androstenedione and testosterone (241). The main urinary metabolite of 17-OHP, pregnanetriol, can also be used to diagnose 21-hydroxylase deficiency. Moreover, urinary levels of pregnanetriol glucuronide may be a way to monitor therapeutic efficacy and possible overtreatment (242). As an alternative to enzyme-linked immunoassays or RIAs, urinary steroid metabolites can be analyzed by GS/MS, in which case several relevant markers for CAH and other disorders of steroid metabolism can be assayed simultaneously (243).

V. Treatment

A. Glucocorticoid replacement

1. Overview. All patients with classic 21-hydroxylase deficiency, and symptomatic patients with nonclassic disease, are treated with glucocorticoids. This suppresses the excessive secretion of CRH and ACTH by the hypothalamus and pituitary and reduces the abnormal blood levels of adrenal sex steroids. In children, the preferred cortisol replacement is hydrocortisone (i.e., cortisol itself) in doses of 10 to 20 mg/M2/day in two or three divided doses. These doses exceed physiological levels of cortisol secretion, which are 6–7 mg/M2/day in children and adolescents (244, 245). Although cortisol secretion is normally only slightly higher in neonates—7–9 mg/M2/day (246)—infants with CAH are usually given a minimum of 6 mg/day in three divided doses. The supraphysiological doses given to children with CAH seem to be required to adequately suppress adrenal androgens and to minimize the possibility of developing adrenal insufficiency.

The short half-life of hydrocortisone minimizes growth suppression and other adverse side effects of longer acting, more potent glucocorticoids such as prednisone and dexamethasone. On the other hand, a single daily dose of a short-acting glucocorticoid is ineffective in controlling adrenocortical hormone secretion (247).

Cortisone acetate is not a drug of first choice for CAH. It has only 80% of the bioavailability of hydrocortisone and approximately two thirds of its potency (248). Moreover, since cortisone must be converted to cortisol to be biologically active, defective 11β-hydroxysteroid dehydrogenase reductase activity can further reduce the efficacy of this drug (249).

Older adolescents and adults may be treated with modest doses of prednisone (e.g., 5–7.5 mg daily in two divided doses) or dexamethasone (total 0.25–0.5 mg given as one or two daily doses). Patients should be monitored carefully for signs of iatrogenic Cushing’s syndrome such as rapid weight gain, hypertension, pigmented striae, and osteopenia. Men with testicular adrenal rests may require higher dexamethasone doses to suppress ACTH.

Treatment efficacy (i.e., suppression of adrenal hormones) is assessed by monitoring 17-OHP and androstenedione levels. Testosterone can also be a useful parameter in females and prepubertal males. Because of the adverse effects of overtreatment (see the next section) it is not desirable to completely suppress endogenous adrenal corticosteroid secretion. A target 17-OHP range might be 100-1000 ng/dl with commensurate age and gender-appropriate androgen levels (247, 250). Hormones should be measured at a consistent time in relation to medication dosing, preferably at 0800 h at the physiological peak of ACTH secretion, or at least at the nadir of hydrocortisone blood levels just before the next dose.
is to be given. Remote monitoring of hormonal control in CAH patients is possible through the use of either salivary hormone measurements (194, 251, 252) or finger-prick blood samples collected on filter paper and assayed for 17-OHP (247, 250, 253). The latter methodology is routinely used for neonatal screening for CAH (see Section IV.B).

Children should have an annual bone age x-ray and careful monitoring of linear growth. Despite careful monitoring and good patient compliance, most retrospective reviews (94–99) indicate that final height averages 1–2 SDs below the population mean or the target height based on parental heights.

2. Adverse effects of overtreatment. Early excessive glucocorticoid treatment (hydrocortisone dose > 20 mg/m²/day) is potentially detrimental to growth (67). A randomized prospective crossover trial showed that patients treated with 15 mg/m²/day of hydrocortisone were less likely to show growth suppression compared with those taking doses of 25 mg/m²/day (254). High body mass index in childhood also correlates with poor final height and may be a surrogate marker for overtreatment (67, 255). However, patients with CAH may be more prone to obesity than other children, and they begin gaining weight earlier in childhood (nadir for adiposity of 1.7 yr in British children with CAH as compared with 5.5 yr in the general UK population) even when height is normal (256). Despite linear growth averaging approximately 1 SD below the mean, bone mineral density does not appear to be compromised in CAH patients receiving typical glucocorticoid doses (98, 257–259). Only one study of Finnish patients showed low bone density in the femoral and L2–4 vertebrae. The authors attributed their findings to excessive glucocorticoid dosing in some subjects (260). However, decreased bone turnover has also been associated with CAH (261).

If control cannot be achieved with hydrocortisone, it is reasonable to use either prednisone or dexamethasone for a 2- to 4-day course of suppressive therapy before resuming hydrocortisone. After epiphyseal fusion, prednisone or dexamethasone may be used as maintenance therapy, but doses should not exceed the equivalent of 20 mg/m² hydrocortisone daily, and patients should be carefully monitored for signs of iatrogenic Cushing’s syndrome.

3. Stress dosing. Patients with classic CAH cannot mount a sufficient cortisol response to stress and require pharmacological doses of hydrocortisone in such situations as febrile illness and surgery under general anesthesia. Such treatment should approximate typical endogenous adrenal secretion in critically ill and perioperative patients (71). Dose guidelines include tripling the maintenance dose of oral hydrocortisone (administered in three divided doses) in minor febrile illnesses. If a patient is unable to tolerate oral medication, intramuscular hydrocortisone sodium succinate (Solu-Cortef) may be given, but medical advice concerning the need for intravenous hydration should be promptly sought. Patients and parents should receive instructions for these types of emergency contingencies, and patients should carry or wear identification with information about their medical condition. For major surgery, administration of hydrocortisone (100 mg/m²/day) divided in four intravenous doses is warranted for at least 24 h peri- and postoperatively before tapering over several days to a maintenance dose. Intravenous hydrocortisone is preferred over equivalent glucocorticoid doses of methylprednisolone (Solu-Medrol) or dexamethasone because (when it is administered in high doses) its mineralocorticoid activity is able to substitute for oral fludrocortisone.

Patients with nonclassic CAH do not require stress doses of hydrocortisone for surgery unless they have iatrogenically been rendered hypoadrenal by prior chronic administration of glucocorticoids. In our experience no patient with nonclassic CAH has ever shown evidence of adrenal insufficiency. However, one assumes that all patients treated over long periods of time with glucocorticoids have some degree of endogenous adrenal suppression. It is therefore prudent to treat with supplemental glucocorticoids in times of extreme stress, and patients receiving such therapy should wear medical alert tags. Alternatively, if given adequate advance notice, one could discontinue treatment and test the integrity of the hypothalamic-pituitary-adrenal axis with a low-dose ACTH stimulation test (234, 235).

4. Indications for therapy in patients with nonclassic CAH. Individuals diagnosed with nonclassic CAH should be offered treatment when they manifest signs or symptoms of androgen excess. Low-dose glucocorticoid therapy may be initiated in children with precocious pubarche, i.e., inappropriately early onset of body hair and odor, accompanied by advanced bone age. A small group of Jewish nonclassic CAH patients was able to achieve final heights within the range predicted from parental heights as long as glucocorticoid therapy was started at the first signs of precocious adrenarche or bone age acceleration; delaying initiation of therapy until after central puberty began was associated with decreased final height (262). Other studies have found no adverse effect of nonclassic CAH on height (187).

Other common indications for treatment are hirsutism, oligomenorrhea, and acne in young women. Infertility patients diagnosed with nonclassic CAH should also be treated, as they may more readily become pregnant if the hormonal imbalance is the principal obstacle to conception. Treatment with glucocorticoids suppresses adrenal androgen production, resulting in gradual improvement in clinical signs of androgen excess. Remission of hirsutism is the most difficult objective to achieve with glucocorticoid monotherapy, as established hair follicles are difficult to eradicate. Cosmetic therapy is therefore advised as an adjunct to hormonal therapy in women for whom the hirsutism is unsightly. An exact timetable to regression of each clinical sign has yet to be established.

Men with nonclassic CAH may achieve improved sperm counts and fertility with glucocorticoid treatment (263, 264). Although rare, testicular enlargement in nonclassic males is also an indication for glucocorticoid therapy (183).

Nonclassic patients whose symptoms have resolved (e.g., a boy treated for precocious pubarche, now fully grown), or affected women past child-bearing age, should be given the option of discontinuing therapy.
B. Mineralocorticoid replacement

Infants with the salt wasting form of 21-hydroxylase deficiency require mineralocorticoid (fludrocortisone, usually 0.1–0.2 mg but occasionally up to 0.4 mg daily) and sodium chloride supplements (1 to 2 g daily; each gram of sodium chloride contains 17 mEq of sodium) in addition to glucocorticoid treatment. The sodium content of either breast milk or the most popular infant formulae is about 8 mEq/liter, which is only sufficient for maintenance sodium requirements in healthy infants. Considerably more sodium (~8 mEq/kg/day) must be supplied to keep up with ongoing losses in aldosterone-deficient infants. Often, older children acquire a taste for salty food and do not require daily supplements of sodium chloride tablets. Moreover, fludrocortisone doses may often be decreased after early infancy.

Although patients with the simple virilizing form of the disease by definition secrete adequate amounts of aldosterone, they are nevertheless often treated with fludrocortisone. This can aid in adrenocortical suppression, reducing the dose of glucocorticoid required to maintain acceptable 17-OHP levels (236).

PRA may be used to monitor mineralocorticoid and sodium replacement. Hypertension, tachycardia, and suppressed PRA are clinical signs of overtreatment with mineralocorticoids (265). Excessive increases in fludrocortisone dosage may also retard growth (266).

C. Other therapeutic approaches

1. Pharmacological. A novel four-drug regimen for CAH, consisting of flutamide (an androgen receptor-blocking drug), testosterone (an aromatase inhibitor), low-dose hydrocortisone, and fludrocortisone, showed promising results after a 6-month trial. Children in the experimental treatment group showed less bone age advancement and more appropriate linear growth velocity than those in the standard treatment group (267). After 2 yr, the 16 children in the experimental group showed higher levels of 17-OHP, androstenedione, DHEA and its sulfate, and testosterone, plus a slower rate of growth and bone maturation with improved predicted height compared with children on standard therapy. However, central precocious puberty occurred and required treatment with LHRH analog in 3 of 8 males in the experimental therapy group and in 0 of 9 control males (268). It remains to be seen whether longer term, larger scale studies will show a favorable effect of the experimental regimen on final height. Other questions include whether the average family could cope with such a medical regimen in a school-aged child, and what the cost of such therapy would be over many years.

Another interesting experimental CAH therapy is the addition of carbenoxolone, an inhibitor of 11β-hydroxysteroid dehydrogenase (11-HSD). The latter is an enzyme important in inactivating cortisol and preventing its access to the mineralocorticoid receptor (269). The rationale for carbenoxolone as an adjunct to therapy of CAH is that inhibition of the oxidative 11-HSD reaction should generate higher endogenous bioactive cortisol levels without administering larger doses of steroids. In a short-term pilot study with an open-label, crossover design involving six CAH patients aged 15 to 39 yr, there were significant reductions in 17-OHP, androstenedione, renin, and urinary pregnanetriol when carbenoxolone was added to the standard therapeutic regimen (270, 271). Hypertension is potentially a complication of such a regimen (269).

Two adult patients with CAH and concurrent malignancies were treated with chlormadinone acetate, an antiandrogen used overseas for prostate cancer. In both cases, secretion of ACTH and adrenal androgens was suppressed (272). At present, antiandrogens are not recommended for treatment of children and young women with CAH outside the research setting, since the risks of adverse side effects, including hepatic toxicity and teratogenicity, are significant.

2. Adrenalectomy. Consequences of inadequate treatment or noncompliance for the female include ongoing virilization in addition to compromise of linear growth. For this reason, it has been suggested that (laparoscopic) adrenalectomy may represent an alternative to suppressive medical therapy with glucocorticoids (273). Severely affected patients, especially females, could perhaps be more easily managed as Addisonians with low-dose glucocorticoids and mineralocorticoids than with adrenal glands that secrete excessive sex steroids. Opponents of surgical treatment feel that this is too radical a step, potentially placing patients at risk from the surgical procedure, and later incurring further risks from iatrogenic adrenal insufficiency. Moreover, the beneficial effects of adrenalectomy may be confounded by the development of gonadal adrenal rests that can secrete androgen precursors (172).

Finally, there is a clear benefit in terms of improved lipid profile, libido, and quality of life from physiological adrenal DHEA production (274, 275) that would be lost with adrenalectomy. Although patients have been managed in this manner (172, 276), further data must be collected before deciding whether adrenalectomy is a viable therapeutic alternative. It is likely to be used, if at all, in patients with severe 21-hydroxylase deficiency refractory to standard medical management.

3. Gene therapy. Because 21-hydroxylase deficiency is an inherited metabolic defect, the question arises of the feasibility of gene therapy (277). Indeed, mice with 21-hydroxylase deficiency refractory to standard medical therapy, albeit not perfect, is effective and relatively inexpensive. High level expression would need to be targeted to the adrenal cortex, where adequate levels of steroid precursors are available. As the most difficult therapeutic goal to achieve is adequate suppression of adrenal androgens, expression would need to be sufficiently high to permit nearly normal levels of cortisol biosynthesis under both normal and stress conditions, and such levels of expression would need to be maintained indefinitely to be cost effective in comparison with conventional treatment. These criteria seem unlikely to be met for the foreseeable future.

D. Corrective surgery

The general approach to evaluating the newborn with ambiguous genitalia has been discussed in Section IV.A. In
general, the recommended sex assignment should be that of the genetic/gonadal sex, if for no other reason than to retain the possibility of reproductive function. This is especially true for females with 21-hydroxylase deficiency who have normal internal genital structures and potential for childbearing. An exception to this rule might be the genetically female patient with completely male appearing genitalia, especially if the child has been raised as a male for more than a few months. Such children will need to be castrated at puberty to avoid feminization.

Whether, how, and when to intervene surgically in the correction of genital anomalies is the subject of continuing debate (279, 280). Some adult patients with CAH and other intersex conditions who are unhappy with their gender assignment, as well as some physicians, have advocated postponing genital surgery until the affected individual is able to provide informed consent for cosmetic genital surgery, and select the gender with which he/she will be most comfortable (279, 281–283). It is not clear, however, whether families would readily accept the idea of raising a child with indeterminate gender and/or ambiguous genitalia, whether children would then be psychologically traumatized due to lack of societal acceptance of such conditions, and whether such children would be able to develop an unambiguous gender identity at all.

It must also be recognized that recommendations for sex assignment are to some extent culture specific. In cultures that value infant boys over girls, parents may strongly resist rearing a female with ambiguous genitalia as a girl, and many girls with severely virilized external genitalia will be raised as males (152, 153).

The most common current approach to surgical correction is for clitoroplasty (284, 285), rather than clitoridectomy, to be done in infancy. In adolescence the patient can be taught to perform vaginal dilation with acrylic molds (286, 287). Vaginal reconstruction is often postponed until the age of expected sexual activity (288, 289), but single-stage corrective surgery has also been successfully performed in children (284, 290, 291). Correction in infancy may be more successful for cases of simple labial fusion than in cases where the distal vagina must be reconstructed (289, 292). Newer modifications in vaginoplasty procedures may improve outcome in patients with urogenital sinus for whom simple dilation is not helpful (286, 293, 294). According to self-assessment surveys among sexually active women with CAH, approximately 60% are able to have satisfactory intercourse (295). Reoperation is frequently required to achieve satisfactory results (292).

As surgical and medical treatment regimens have improved in recent years, more women with CAH have successfully conceived spontaneously, completed pregnancies, and given birth (296). Most often delivery is by cesarean section due to an inadequate introitus, but vaginal delivery is possible in some cases (109).

E. Psychological counseling

Families of CAH patients should be assessed for emotional health. The initial screening will most likely be done by the pediatrician and pediatric endocrinologist. The child behav-ior checklist and the self-perception profile can be used in school-aged children (140). Parents should be offered psychological counseling soon after the diagnosis is made. Intermittent assessment of family functioning, as has been done in other disease states, may be a useful tool in predicting future problems (297). Children should, as they mature, be repeatedly informed about their condition by parents and physicians in a sensitive and age-appropriate manner. When psychotherapy is undertaken, medical and psychiatric caregivers should maintain communication so that both are aware of the patient’s and family’s status. Unfortunately, many locales lack mental health professionals with experience in counseling patients and families with intersex conditions.

Although psychosexual development of females with classic CAH is incompletely understood (Section III.H), we believe anticipatory counseling of patients’ families should initially address the high likelihood that affected girls will exhibit tomboyish behavior, masculine play preferences and perhaps, when older, a preference for a career over domestic activities (133). In the contemporary United States, these preferences usually have a high degree of social acceptance, considering the increased availability of and interest in girls’ competitive sports as well as the many women who work. Parents should also be reassured that the majority of (but not all) girls function heterosexually, although they may require repeated genital surgeries to have satisfactory intercourse. The endocrinologist and/or mental health professional (depending on inclination and experience) caring for the adolescent girl with CAH should address sexual orientation, both fantasized and actual. The patient should be reassured that some degree of attraction to other girls, although it does not always occur, is a typical feature of her condition. A discussion of psychotherapy for homosexuality is beyond the scope of this review, but it should be accepted by health care professionals that a minority of women with CAH may be most comfortable as homosexuals and that such individuals should be helped to come to terms with their situation. Adult patients should also be made aware of relevant patient advocacy groups.

F. Treatment of precocious puberty

Central precocious puberty may occur in the setting of excess adrenal sex steroids and advanced bone age, especially when glucocorticoid treatment is initiated in children with markedly advanced bone age. Under such circumstances, chronic exposure to adrenal androgens may cause the hypothalamic-pituitary gonadal axis to mature. A sudden decrease in androgen levels with adequate treatment may then trigger secretion of gonadotropins by the pituitary.

Clinical suspicion of central precocious puberty in affected boys may be aroused when physical examination in boys reveals increased testicular growth, or when girls show increased breast growth. In boys, serum testosterone may increase in the face of well controlled 17-OHP. However, since testosterone is also a byproduct of excessive adrenal sex hormone production, this hormone alone is not an accurate marker for central precocious puberty. An elevated ratio of testosterone to androstenedione suggests a gonadal, rather
than an adrenal, source of hormone production. Serum estrogen levels are not typically part of the hormonal profile for CAH management.

Definitive diagnosis of precocious puberty requires GnRH stimulation testing. LH and FSH levels drawn before and 30 min after a 100-μg bolus of GnRH (Factrel) will show a marked rise in LH > FSH; the absolute levels depend on the type of assay employed.

Although spontaneous resolution of central precocious puberty has been anecdotally reported (298), this condition usually requires separate pituitary suppressive treatment with GnRH analogs (299). The goals of GnRH analog treatment are to suppress pituitary gonadotropins and, consequently, gonadal sex steroid production and to attempt to enhance adult height by preventing premature epiphyseal fusion. Whereas the first goal is readily achieved, the latter objective is more difficult. Preliminary data consisting solely of predicted height in but a few children allows cautious optimism about such treatment, but more long-term data are needed (300). Another goal of therapy is to avoid the adverse psychological consequences of premature puberty; anecdotal experience suggests this is aided by suppressive medical therapy.

GH-deficient children with precocious puberty who do not have CAH appear to benefit from combined treatment with GH and GnRH agonists if such treatment is begun at a relatively young bone age (301). However, short but otherwise healthy children with normally timed puberty do not benefit in terms of final height outcome from GnRH agonist plus GH therapy (302). CAH children are typically not GH deficient. Thus, multidrug regimens such as these remain experimental and very costly, and their effectiveness has not been established in CAH.

G. Prenatal therapy

1. Overview. In pregnancies at risk for a child affected with virilizing adrenal hyperplasia, suppression of fetal adrenal androgen production and decreased genital ambiguity in females have been achieved by administering dexamethasone to the mother (303–311). As compared with hydrocortisone, dexamethasone has no salt retaining activity and it is able to cross the placenta because it is not metabolized significantly by placental 11β-hydroxysteroid dehydrogenase (269). The dose is typically 20 μg/kg/day based on prepregnancy weight to a maximum of 1.5 mg daily in three divided doses, beginning before the 7th to 8th week of gestation (307, 311). Approximately 70% of prenatally treated females are born with normal or only slightly virilized genitalia. Treatment failures, i.e., affected females requiring genital reconstruction, have been attributed to late onset of treatment, cessation of therapy in midgestation, noncompliance, or suboptimal dosing (312), whereas others had no ready explanation (313). The fetal adrenal cortex may not always be adequately suppressed by these doses of maternally administered dexamethasone. Unaffected newborns treated until birth usually have suppressed steroid secretion for at least 1 week after birth, especially of steroids such as 16α-hydroxyprogrenolone that originate in the fetal zone of the cortex. However, 17-OHP and 21-deoxycortisol metabolites in affected infants may not be suppressed at birth even with continuous treatment beginning early in gestation (314).

Very few data are available regarding natural variability in genital virilization among family members with the same CYP21 genotype, although nearly normal genitalia without treatment in relatives of highly virilized patients have been reported anecdotally (315, 316). Nevertheless, the likelihood of severe genital ambiguity is apparently reduced among those prenatally treated compared with their untreated sisters and to all affected girls with similar genotypes. Anecdotal reports suggest that suppression of adrenal androgen secretion is easier to achieve after prenatal treatment (310). The first prenatally treated female has reached late adolescence with normal cognitive development (310). It is not yet possible to determine whether prenatal treatment will induce marked differences in the psychosexual outcome for women with CAH.

Prenatal therapy is usually coupled with prenatal diagnosis (see Section VI.H). Since dexamethasone treatment suppresses amniotic fluid adrenocortical hormones, genetic diagnosis must be performed. Although the incidence of fetal deaths in treated pregnancies does not appear to exceed that for the general population [9% spontaneous abortion rate in treated pregnancies, compared with 14% in untreated pregnancies (317)], a high rate of spontaneous abortions has been observed after chorionic villus sampling (CVS) performed to obtain tissue for genetic diagnosis (318). Either amniocentesis or CVS may be done for diagnostic purposes, but the latter should be carried out only in experienced centers at 10–12 weeks gestation (319). If the sex is male, or CYP21 genotype indicates the fetus is unaffected, dexamethasone should promptly be discontinued to minimize potential risks of glucocorticoid toxicity (Fig. 9).

2. Therapeutic risks to the fetus. Most published literature dealing with fetal exposure to glucocorticoids deals with late second or third trimester treatment. Prenatal treatment of CAH is different in that it must begin in the early first trimester to be effective in preventing female genital ambiguity. Since only one of eight fetuses is likely to be an affected female, seven of eight pregnancies will be unnecessarily exposed to at least several weeks of dexamethasone treatment.

![Fig. 9. Flowchart for decisions pertaining to prenatal diagnosis of 21-hydroxylase deficiency. Format is identical to Fig. 5.](https://academic.oup.com/edrv/article-abstract/21/3/245/2423831)
before the sex and disease status can be determined. Because the unaffected offspring will not derive any benefit from glucocorticoid treatment, the issue of potential risk is all the more significant. Normal fetal cortisol production is considerably lower than that of children or adults; samples obtained during fetoscopy, show umbilical vein cortisol at 16–20 weeks gestation to average approximately 20 nmol/liter (0.72 μg/dl) (320). Thus, providing pharmacological doses of a potent, long-acting glucocorticoid such as dexamethasone might disrupt fetal physiology. Potential risks include congenital malformations such as cardiac septal hypertrophy (321), hydrometrocolpos (322), and hydrocephalus (318). Moreover, the effects of abrupt dexamethasone withdrawal on fetal development are unknown. Intrauterine growth retardation and unexplained fetal death have been observed in 2% or less of treated pregnancies (317); these statistics are not significantly different from those found in the population at large (323). The risk of overt human fetal defects appears to be low compared with complications observed in a rodent model of in utero exposure to high-dose glucocorticoids (324), which features frequent cleft palate in addition to fetal growth retardation and/or demise. Pregnant rats treated with dexamethasone, 20 μg/kg/day (the same weight-based dose used in human prenatal treatment for CAH), produced litters with average birth weights 14% below those of controls; the offspring were also hypertensive at 6 months, i.e., young adulthood (325). Mid-to-late gestation administration of dexamethasone to fetal rats alters levels of transcription factors important in brain development (326). It is conceivable that more subtle effects of glucocorticoids on the developing human brain may go unnoticed during early life. A pilot study using maternal surveys suggested that children not affected with CAH who were subjected to prenatal dexamethasone treatment were more shy than untreated children (327). However, prenatally treated children have not undergone thorough neuropsychological testing. Recent editorials called attention to these issues, citing numerous studies in animal models showing the dangers of prenatal exposure to glucocorticoids with respect to impairment of somatic growth, brain development, and blood pressure regulation (328–330). An international registry might facilitate long-term studies that could answer many of these questions.

3. Therapeutic risks to the mother. The incidence of maternal complications has varied among investigators; overall, it is about 10% (310). In evaluating such data, it is important to correlate the types of adverse side effects observed and the duration of therapy. Both American (331) and European (331) investigators have found a higher incidence of side effects in women treated from the first through third trimesters. Cushingoid features, excessive weight gain, severe striae, hypertension, and hyperglycemia are seen in this setting. These side effects most often resolve with discontinuation of treatment. Weight, blood pressure, and glucose tolerance should be closely monitored in all treated women treated to term. Serial maternal serum dexamethasone levels, if available, might prevent over- and undertreatment, but they have not been followed routinely. Maternal urinary estriol measurements have also been suggested as a guide to adjusting maternal treatment (304). A gradual decrease in the dose of dexamethasone later in gestation might decrease the incidence of maternal side effects, but there are as yet no data concerning the efficacy of such a regimen. More common maternal side effects in those treated for a shorter duration include edema, gastrointestinal upset, mood fluctuations, acne, and hirsutism; one or more of these symptoms are seen in 10–20% of women treated in early pregnancy.

Because of these concerns, caution should be exercised in recommending prenatal therapy with dexamethasone, and women must be fully informed of potential fetal and maternal risks, some of which may be as yet unrecognized. Additionally, the possibility of lack of therapeutic benefit should be disclosed when obtaining informed consent. However, we emphatically disagree with a recent suggestion (330) that prenatal treatment is so experimental as to require approval by institutional review boards.

Caveats notwithstanding, many parents of affected girls still opt for prenatal medical treatment because of the severe psychological impact of ambiguous genitalia on the child and on the family (332). Similar diagnostic and therapeutic approaches can also be effective in families at risk for 11β-hydroxylase deficiency, in which affected female fetuses may also suffer severe prenatal virilization (333).

4. Other considerations for genetic counseling. The most common circumstance in which prenatal treatment is offered to a pregnant woman is when she and her partner have already had a child with CAH, in which case the likelihood of her bearing an affected girl with each successive pregnancy is 1/8.

Other scenarios may arise in genetic counseling. What if one partner has classic CAH and the carrier status of the partner is not known (recognizing that an affected mother will remain on glucocorticoid replacement regardless, but her dose may need to be increased to treat the fetus)? If the carrier frequency for classic CAH in the general population is 1.6% (equivalent to a disease frequency of 1/16,000, see Section IV.B), then the a priori likelihood of these parents giving birth to an affected girl is 0.4%, or 1/250 ([1 parent carrying 2 classic alleles] × [1.6% carrier frequency in general population] × [½ chance the carrier parent will pass his or her affected allele to the fetus] × [½ chance the fetus is female]).

Infants of mothers affected with the nonclassic disorder are also at slightly increased risk of developing classic CAH. Most studies examining genotypes in CAH (Section VI.I) did not specifically target nonclassic patients, and ascertainment biases and ethnic differences make it difficult to draw firm conclusions regarding allele frequencies in the nonclassic CAH patient population. Nevertheless, it appears that at least 50% of women clinically ascertained to have nonclassic CAH are compound heterozygotes for a classic and a nonclassic mutation (209, 210, 334). Accepting this figure, there is a priori approximately a 0.1% (1/1000) chance that a mother with nonclassic disease will give birth to a daughter affected with classic CAH ([50% carrier frequency of classic alleles among women with nonclassic disease] × [1.6% carrier frequency in general population] × [¼ chance both classic alleles will be passed to the fetus] × [½ chance the fetus is female]).

In each of these scenarios, the risk of having an affected
daughter is far less than the 1:8 risk with two known classic 21-hydroxylase deficiency carriers. Thus, prenatal therapy is not warranted unless the carrier status of the mate (as well as the genotype of a nonclassic patient) is first ascertained by hormonal testing and/or genotyping as part of preconception genetic counseling. This may be desired by some but not all couples. It may nevertheless be prudent to measure 17-OHP in such at-risk infants after birth, remembering that basal hormone measurements may not be sufficient to diagnose nonclassic CAH in infants and young children.

VI. Molecular Genetic Analysis

A. Biochemistry of CYP21

Steroid 21-hydroxylase (P450c21, CYP21) is a microsomal cytochrome P450 enzyme that converts 17-OHP to 11-deoxy-cortisol and progesterone to DOC (335–341). As with other microsomal P450s, the enzyme accepts electrons from an NADPH-dependent cytochrome P450 reductase (336), thus reducing molecular oxygen and hydroxylating the substrate. The P450 reductase is required because NADPH donates electrons in pairs, whereas P450s can only accept single electrons (342).

Human CYP21 normally contains 494 amino acid residues (343, 344) [a normal variant has an extra leucine within the N-terminal hydrophobic domain and thus contains 495 residues (345)] and has a molecular mass of approximately 52 kDa. When expressed in mammalian cells, human recombinant CYP21 has apparent K_m values for 17-OHP and progesterone of 1.2 and 2.8 μM, respectively, and the apparent V_max for 17-OHP is twice that of progesterone (346).

B. Structure-function relationships

Alignment of the amino acid sequences of many P450s have identified a small number of strongly conserved residues that are presumed to be important for catalytic function (347, 348). The basic three-dimensional structure of P450 enzymes has been deduced from x-ray crystallographic studies of four bacterial P450s. The first of these, P450cam (CYP101, camphor 5-exo-hydroxylase from Pseudomonas putida), is a soluble molecule that bears little similarity in primary structure (roughly 15%) to eukaryotic P450s (349); P450terp (350) and P450eryF (351), bacterial P450s structurally related to P450cam, have also been characterized. In contrast, P450BM-3 (CYP102 from Bacillus megaterium), is a complex protein consisting of a P450-
like N-terminal domain and a C-terminal domain that is
35% identical to eukaryotic cytochrome P450 reductase.
The P450 domain is 25–30% identical to the eukaryotic
CYP4 and CYP52 families (352) and approximately 20%
identical to CYP21 (Fig. 10). This domain has been sub-
ject to crystallographic analysis (353). Its sequence has
been aligned with CYP21 and used as the basic for three-
dimensional modeling of CYP21 (354). Based on thermo-
dynamic considerations, this model may not be as accurate
as analogous models of other eukaryotic P450s such as
CYP19 (aromatase) (355) or CYP17 (17-hydroxylase) (356).
Nevertheless, based on these analyses and functional stud-
ies, several conclusions may be drawn.

1. Heme binding. The heme iron is critical for catalytic function
of P450s. Of its six coordination positions, four interact with
the protoporphyrin ring. One of the two axial positions is
coordinated to the sulphydryl group of a completely con-
erved cysteine (C428 in CYP21), which is located in a rel-
atively highly conserved “heme binding peptide” near the
C-terminus. Mutation of C428 in CYP21 destroys enzymatic
activity (357).

2. Oxygen and water binding. The ligand at the other axial
position of heme is either a water or an oxygen molecule.
When an oxygen molecule is bound, it is parallel to the axis
of coordination with the iron atom. An H2O molecule is
consistently present in a groove in the “I” helix adjacent to
strongly conserved acidic (aspartate or glutamate) and thre-
onine residues (E294 and T295 in CYP21). Mutation of the
conserving threonine in several other P450s destroys or
Drastically decreases enzymatic activity (358, 359).

According to one of the several proposed models of P450
catalysis (360), the first step of the reaction is binding of
substrate to oxidized (ferric, Fe3+) enzyme. One electron is
donated from P450 reductase to the enzyme so that the iron
is in the reduced (ferrous, Fe2+) state. This complex binds
molecular oxygen and then accepts a second electron from
the accessory protein, leaving the bound oxygen molecule
with a negative charge. Two protons are then donated in
succession to the water molecule by the carboxyl group of the
acidic residue, transferred to the hydroxyl of the conserved
threonine, and finally donated to the distal oxygen atom
(353). The distal oxygen atom is then released as a water
molecule, leaving the iron in the Fe2+ state. The remaining
oxygen atom is highly reactive (the iron-oxygen complex is
a “ferryl” moiety) and attacks the substrate, resulting in an
hydroxylation.

3. Substrate binding. Like most P450 substrates, steroids are
relatively hydrophobic molecules. Thus, it is likely that the
substrate binding site(s) will consist primarily of hydropho-
bic amino acid residues.

A priori, it was possible that the substrate binding sites of
steroid-metabolizing P450s would more closely resemble
each other in sequence than the substrate binding sites of
other P450s such as xenobiotic metabolizing enzymes. Com-
parisons of the sequences of 21 and 17-hydroxylase and
cholesterol desmolase (CYP21, CYP17, and CYP11A) iden-
tified two highly conserved areas, one near the N terminus
(Q53-R60 in CYP21) and the other toward the C terminus
(L342-V358 in CYP21) (361, 362).

The crystal structure of CYP102 confirms that part of the
first of these indeed corresponds to a portion of a deep pocket
constituting the substrate binding site (β-sheet 1–1, residues
E38-A44 in CYP102). However, the second conserved area
corresponds to helix K (L311-W325 of CYP102), which does
not interact with substrate. Instead, this region forms part of
the docking site for the accessory electron transport protein,
cytochrome P450 reductase. The remainder of segments that
form the substrate binding pocket are widely distributed in
the peptide (remainder of β-sheet 1, B’ and F helices) and the
sequence conservation among steroid metabolizing P450s is
not particularly strong in these regions.

4. Binding to accessory proteins. Microsomal and mitochondrial
P450s accept electrons from cytochrome P450 reductase or
adrenodoxin, respectively. In either case chemical modifica-
tion studies suggest that basic amino acids (usually lysine) on
the P450 interact with acidic residues on the accessory pro-
tein. The crystallographic studies of CYP102 suggest a dock-
ing site for reductase formed in part by helices B, C, D, J’, and
K. Helix K, as mentioned, was previously thought to interact
with substrate. Support for the idea that it is instead required
for redox interactions (with cytochrome P450 reductase or
adrenodoxin, depending on the type of P450) comes from
mutagenesis studies of CYP11A, wherein modification of
either of two lysine residues in helix K destroys enzymatic
activity without affecting substrate binding (363). In similar
studies of CYP17, mutagenesis of arginine residues in this
region disrupts interactions with cytochrome P450 reductase
cytochrome b5 (364). There are several naturally occurring
mutations of arginine residues (R354 and R356) in this
region of CYP21 that drastically decrease enzymatic activity
(365–367).

Only two basic residues, one in helix K and the other in the
heme binding peptide (R323 and R398 in CYP102, corre-
sponding to R354 and R426 in CYP21), are completely con-
served in all eukaryotic P450s, suggesting that other posi-
tively charged residues that are not completely conserved
may be necessary for binding to the accessory protein (368).

C. CYP21 gene structure

The structural gene encoding human CYP21 (CYP21,
CYP21A2, or CYP21B) and a pseudogene (CYP21IP, CYP21A1P,
or CYP21A) are located in the HLA major histocompatibility
complex on chromosome 6p21.3 approximately 30 kb apart,
adjacent to and alternating with the C4B and C4A genes en-
coding the fourth component of serum complement (369, 370)
(Fig. 11). In addition the RPI (G11) gene is located immedi-
ately 5’ of C4A and encodes a putative nuclear protein similar
to DNA helicase; a truncated copy of this gene, RP2, is
located between CYP21P and C4B (371, 372). The CYP21, C4, and RP
genes are transcribed in the same direction. CYP21 overlaps a
gene on the opposite DNA strand (OSG, XB, or TNXB) that
encodes a putative extracellular matrix protein, tenasin-X
(373). CYP21P overlaps a truncated copy of this gene (TNXA)
that does not encode a functional protein.

CYP21 and CYP21P each contain 10 exons spaced over 3.1
D. Transcription

1. Naturally occurring transcripts. For purposes of this review, the most important gene transcript in the C4-CYP21 region is that of CYP21 itself, which begins 10–11 nt before the initial AUG codon. Whereas the C4A, C4B, and TNXB genes are mainly expressed in other tissues, the truncated TNXA gene is transcribed in an adrenal-specific manner (374).

CYP21P is also transcribed specifically in the intact adrenal cortex at a level 10–20% that of CYP21(375). However, the first 2 introns are inconsistently spliced out, and an uncertain proportion of transcripts include additional exons in the region between the end of CYP21P and the beginning of C4B. Some of these exons may overlap the truncated TNXA gene (375). In contrast, CYP21P transcripts cannot be detected in primary cultures of human adrenocortical cells, whereas CYP21 is appropriately expressed under the same conditions (376, 377). In any case, CYP21P transcripts do not contain a long open reading frame and are of uncertain functional significance (378). Adrenal transcripts in the same direction as CYP21 have also been detected overlapping TNXB; these are also of uncertain functional significance (375, 378).

2. Hormonally induced expression. The primary factor regulating CYP21 expression in the zona fasciculata of the adrenal cortex is ACTH (reviewed in Ref. 20). ACTH induction is mediated mainly through increased transcription (379, 380), is duplicated by cAMP and related agonists such as forskolin, and requires protein kinase A (381). This is consistent with the known mode of action of ACTH.

Factors inducing expression of CYP21 in H295R human adrenocortical carcinoma cells include cAMP (the second messenger for ACTH) and angiotensin II, which acts primarily through the protein kinase C pathway but also through Ca2+ signaling (382). The cAMP and protein kinase C pathways also induce CYP21 expression in primary cultures of human adrenocortical cells, as do insulin and IGF-I (377).

3. 5'-Flanking sequences controlling transcription. In cultured mouse Y-1 or human H295 adrenocortical cells, the 5'-flanking region of human CYP21 drives basal expression of reporter constructs at levels 2.5–8 times higher than the corresponding region of CYP21P (376, 383, 384). Sequences responsible for this difference have been localized to the first 176 nucleotides (376) although sequences upstream of this region are required for full expression (in this discussion, we will number nucleotides from the start of translation, as different numbering systems have been used by different authors). There are only 4 nucleotide differences between CYP21P and CYP21 in the proximal 176 nucleotides (Fig. 12). It appears that the most important differences are at nucleotide −113, which is a G in CYP21 and an A in CYP21P, and at −126, which is a C in CYP21 and a T in CYP21P (384). The latter polymorphism is in the middle of a binding site for the Sp1 transcription factor from −123 to −129; the CYP21P sequence binds this factor much less well. Moreover, −126C

Fig. 11. A, location of the CYP21 genes within the HLA major histocompatibility complex on chromosome 6p21.3. There are many more genes in this region than are indicated. The direction of the centromere is indicated by the circle at the right end of the line. Numbers denote distances between genes in kilobasepairs (kb). HLA-B is the nearest Class I transplantation antigen gene to CYP21, and HLA-DR is the nearest Class II gene. The region between these classes of genes is termed “Class III.” TNP, Tumor necrosis factor (actually two genes). Tenascin-X gene and TNXB, region around the 21-hydroxylase (CYP21) gene. Arrows denote direction of transcription. CYP21P, 21-Hydroxylase pseudogene; C4A and C4B, genes encoding the fourth component of serum complement; RP1, gene encoding a putative nuclear protein of unknown function; RP2, truncated copy of this gene. TNXB, tenascin-X gene and TNXA, a truncated copy of this gene, are on the opposite chromosomal strand. The 30-kb scale bar is positioned to show the region involved in the tandem duplication.

Fig. 12. Transcriptional regulatory sequences in CYP21. The sequences are numbered from the start of translation so that the +1 position is the first nucleotide of the initial ATG, and the −1 position is the first nucleotide 5′ of this. Putative or actual transcription factor binding sites (discussed in the text) (376) are boxed and marked. Segments known to be required for cAMP regulation (386) are underlined. Nucleotides at which CYP21 and CYP21P differ are denoted by the CYP21P sequence below the CYP21 sequence; those in boldface affect gene expression (384). The triangle denotes start site for transcription.
is at one end of an overlapping binding site for an additional transcription factor termed "adrenal specific protein" (ASP). ASP has not yet been fully characterized but is presumed to bind DNA through zinc fingers as do nuclear hormone receptors (385–387). In contrast, −113G does not lie within a canonical binding site for any known transcription factor, but it is similar to an Sp1 site, and mutation of this nucleotide does interfere with binding of Sp1 (384). There is a site at −110/−103 that could bind the nuclear factor NFB (NF-Mb), which, as far as is known, is specific for granulocytes and macrophages (388). Mutating nucleotide −110 does not have major effects on expression, and so the significance of this putative binding site is uncertain.

Another nuclear factor critical for basal expression of mouse Cyp21 (389) is steroidogenic factor-1 (SF-1, also called Ad4BP). This is an "orphan" nuclear hormone receptor that is required not only for expression of most steroid hydroxylases but also for the embryonic development of the adrenal gland and gonads (reviewed in Ref. 50). Human CYP21 contains a consensus SF-1 site (CCAAGGCCA, with the underlined positions being most important for binding) at −169/−175 (343, 344). However, to the best of our knowledge, the function of this site in the human gene has never been experimentally tested.

Although CYP21 is known to respond to cAMP, the 5′-flanking region does not contain a canonical cAMP response element (a CRE, which is TGACGTCA or a variant thereof). Two regions have been implicated in cAMP responsiveness because they confer such responsiveness on heterologous reporter constructs. The first is a segment from −140/−107, which contains the aforementioned Sp1 and ASP recognition sites (385, 386), and the second extends from −244/−237 and binds the transcription factor NGFI-B (nerve growth factor inducible-B, also called Nur77) (390). Mutation of either of these sites destroys cAMP responsiveness (376, 386); moreover, cotransfection with an NGFI-B expression plasmid transactivates CYP21 reporter constructs (376, 390). NGFI-B is constitutively present in Y-1 mouse adrenocortical cells but is phosphorylated within the DNA-binding domain and does not bind DNA. Treatment with ACTH results in de novo synthesis of unphosphorylated protein that is able to bind DNA and is transcriptionally active (391).

4. More distal elements. Whereas 330 nucleotides of 5′-flanking sequences from the mouse Cyp21 gene are sufficient for expression of reporter constructs in cultured Y-1 cells (381), even 2.2 kb of flanking sequences are unable to direct expression to the adrenal gland in transgenic mice (392). At least 6.4 kb of such sequences are required; the necessary sequences are localized to two short sequences 5–6 kb upstream of Cyp21, located within the adjacent Slp gene (an inactive homolog of complement C4) (392). These sequences may constitute a locus control region. Such regions are required to establish a tissue-specific open chromatin domain in the vicinity of a particular locus and thus permit appropriate tissue-specific expression; the first one identified was in the β-globin cluster (393). In addition to their effects on chromatin, elements within certain locus control regions can act as transcriptional enhancers, and that is the case for both of the elements in Slp when they are tested in mouse Y-1 cells.

Even these sequences may not be sufficient for full expression of Cyp21, because the levels of expression of reporter constructs in transgenic animals are low compared with the intrinsic Cyp21 gene (394).

It is not yet certain whether a similar region exists in humans, but a cryptic adrenal specific promoter has been located within the C4A gene (395). The finding that the truncated TNXA gene is expressed specifically in the adrenal gland (396), although its promoter has been deleted by the duplication of the entire C4-CYP21-TNXA locus, suggests that there is, at least, an adrenal-specific enhancer that is able to influence expression of several adjacent genes.

E. HLA linkage

CAH due to 21-hydroxylase deficiency is inherited as a monogenic autosomal recessive trait closely linked to the HLA complex, meaning that siblings who have 21-hydroxylase deficiency are almost invariably HLA identical (397, 398). Before cloning of CYP21, HLA typing was the main way to perform prenatal diagnosis (399) (see Section VI.H). In addition, particular forms of 21-hydroxylase deficiency are associated with particular combinations of HLA antigens, or haplotypes; this phenomenon is referred to as genetic linkage disequilibrium. The most interesting is an association between the salt wasting form of the disease and HLA-A3; Bw47; DR7 most characteristically seen in Northern European populations. In addition to 21-hydroxylase deficiency, this haplotype usually carries a null allele at one of the two C4 loci encoding the fourth component of serum complement (400, 401). Before cloning of CYP21, this was strongly suspected to represent a contiguous gene syndrome due to a single deletion of C4 and 21-hydroxylase genes; this was confirmed shortly after CYP21 was cloned (2, 370). The deletion apparently occurred after the haplotype was generated, because the identical haplotype without the deletion has been identified in the Old Order Amish (402). The nonclassic form of 21-hydroxylase deficiency is often associated with HLA-B14; DR1, particularly in Eastern European Jewish populations (58, 403, 404). This haplotype is associated with the V281L mutation in CYP21 (see below) and with a duplication of complement C4A and the CYP21P pseudogene (405, 406). Finally, HLA-A1;B8;DR3 is negatively associated with 21-hydroxylase deficiency. This haplotype has a C4A null allele and is associated with deletion of the C4A and CYP21P genes (370, 407). Thus, comparison of a very few individuals homozygous for HLA-A3; Bw47; DR7 or A1; B8; DR3 strongly suggested that the CYP21 gene (then called the 21-hydroxylase "B" gene) was an active gene, whereas the CYP21P gene (21-hydroxylase "A") was a pseudogene (370).

Using pulsed field gel electrophoresis, CYP21 has been mapped approximately 600 kb centromeric of HLA-B and 400 kb telomeric of HLA-DR. It is transcribed in the telomeric to centromeric direction (408, 409).

F. Mutations causing 21-hydroxylase deficiency

Most mutations causing 21-hydroxylase deficiency that have been described thus far are apparently the result of either of two types of recombinations between CYP21, the
normally active gene, and the CYP21P pseudogene. These two mechanisms are unequal crossing over during meiosis, resulting in a complete deletion of C4B and a net deletion of CYP21 (2, 405, 410), or apparent gene conversion events that transfer deleterious mutations normally present in CYP21P to CYP21 (206, 411–417).

The deleterious mutations in CYP21P include an A→G substitution 13 nucleotides (nt) before the end of intron 2 that results in aberrant splicing of pre-mRNA, an 8-nt deletion in exon 3 and a 1-nt insertion in exon 7, each of which shifts the reading frame of translation, and a nonsense mutation in codon 318 of exon 8 (343, 344). There are also 8 missense mutations in CYP21P, 7 of which have been observed in patients with 21-hydroxylase deficiency (Fig. 13 and Table 4).

Because particular mutations occur in many unrelated kindreds, each mutation, and the degree of enzymatic compromise it causes, may be correlated with the different clinical forms of 21-hydroxylase deficiency (i.e., salt wasting, simple virilizing, and nonclassic disease) (Tables 1, 4, and 5, and see Section VI.I) (34, 209, 334, 418–432). The functional effects of missense mutations have been assessed in vitro by recreating them in CYP21 cDNA and expressing the mutant cDNA using an appropriate expression vector. Several systems have been used including transfection of plasmids in mammalian cells (365, 366, 414, 433–437), infection of mammalian cells with recombinant vaccinia virus (346, 438, 439), or expression in yeast (357, 434, 435, 440) or bacteria (441). In general, these systems have yielded similar results regarding the effects of particular mutations on enzymatic activity. The simplest way to compare these studies is to consider whether a particular mutation destroys, drastically reduces, or partially reduces enzymatic activity. A more quantitative way is to express mutant enzymatic activity as a percentage of wild-type activity. For experiments done in whole cells, enzymatic activity is most reliably measured with substrate concentrations below the Km for the enzyme and for incubation times that convert a relatively small proportion of substrate to product (i.e., that remain in the linear phase of the reaction). It is probably meaningless to attempt to derive apparent kinetic constants from measurements in whole cells, considering that glucocorticoids are subject to transport out of cells by mechanisms that may have their own kinetics (442). Conversely, certain CYP21 mutants are relatively unstable when cells are broken (346), and it may be difficult to correlate the apparent kinetic constants obtained under such circumstances with activities in vivo. When apparent kinetic constants are obtained, first-order rate constants (Vmax/Km) are conceptually similar to “enzymatic activity” and tend to be more reproducible across different studies than the individual apparent kinetic constants, Vmax and Km.

1. Deletions and large gene conversions. Large deletions involving C4B and CYP21 comprise approximately 20% of alleles in patients with classic 21-hydroxylase deficiency in most populations (Table 5) but are rarer in some Latin American countries (443, 444). Many deleted alleles are associated with the HLA haplotype A3;Bw47;DR7 (2). Deletions usually extend approximately 30 kb from somewhere between exons 3 and 8 of CYP21P through C4B to the corresponding point in CYP21, yielding a single remaining CYP21 gene in which the 5′-end corresponds to CYP21P, and the 3′-end corresponds to CYP21 (Fig. 14) (396, 410, 445). Deleterious mutations within the CYP21P portion render such a gene incapable of encoding an active enzyme. All patients who carry homozygous deletions suffer from the salt wasting form of the disorder.

Unequal cross-overs may occur anywhere within the duplicated 30-kb region including the RP, C4, and TNX genes (446), and chromosomes with 1 or 3 copies of the 30-kb region occur frequently (447, 448); such rearrangements are found on 16% and 12% of chromosomes 6, respectively (446). Only cross-over breakpoints within or 3′ of the CYP21 genes cause 21-hydroxylase deficiency; breakpoints in the C4 genes delete CYP21P and one of the C4 genes, as is seen with the common HLA haplotype A1;B8;DR3 (370, 407). Thus there is apparently no strong selection against chromosomes with 1 or 3 copies of the 30-kb region as long as CYP21 remains intact.

One kindred carries an unusual deletion extending into the TNXb gene; the patients in this kindred have a contiguous gene syndrome including 21-hydroxylase deficiency and a form of Ehlers-Danlos syndrome caused by loss of function of tenascin-X (449). However, an unrelated patient with 21-hydroxylase deficiency and a similar heterozygous deletion apparently had no additional problems (446).

In most studies (2, 34, 370, 405, 406, 410, 419–421, 423,
Table 4. Mutations causing 21-hydroxylase deficiency

<table>
<thead>
<tr>
<th>Name</th>
<th>Mutation/nt</th>
<th>Ef</th>
<th>Activity % nl</th>
<th>Effect</th>
<th>Clinical severity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td></td>
<td></td>
<td></td>
<td>No enzyme</td>
<td>SW</td>
<td>(2)</td>
</tr>
<tr>
<td>Conversion</td>
<td></td>
<td></td>
<td></td>
<td>No enzyme</td>
<td>SW</td>
<td>(411)</td>
</tr>
<tr>
<td>L9</td>
<td>28 + CTG</td>
<td>E1</td>
<td>100</td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(346)</td>
</tr>
<tr>
<td>W22X</td>
<td>66 G→A</td>
<td>E1</td>
<td>0</td>
<td>Nonsense</td>
<td>SW</td>
<td>(474)</td>
</tr>
<tr>
<td>W22 + Int</td>
<td>64 + T</td>
<td>E1</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(483)</td>
</tr>
<tr>
<td>*P30L</td>
<td>89 C→T</td>
<td>E1</td>
<td>30–60</td>
<td>?Orientation in ER</td>
<td>NC</td>
<td>(438)</td>
</tr>
<tr>
<td>P30Q</td>
<td>89 C→A</td>
<td>E1</td>
<td>0</td>
<td>?Orientation in ER</td>
<td>SW</td>
<td>(469)</td>
</tr>
<tr>
<td>Y471 nt</td>
<td>138 A</td>
<td>E1</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(484)</td>
</tr>
<tr>
<td>Intron 1 splice acceptor</td>
<td>295 A→G</td>
<td>I1</td>
<td></td>
<td>Abnormal splicing</td>
<td>SW</td>
<td>(474)</td>
</tr>
<tr>
<td>G90V</td>
<td>366 G→T</td>
<td>E2</td>
<td>0</td>
<td>Abnormal splicing</td>
<td>SV</td>
<td>(367,475)</td>
</tr>
<tr>
<td>Intron 2 splice donor</td>
<td>387 G→A</td>
<td>I2</td>
<td></td>
<td>Abnormal splicing</td>
<td>SV</td>
<td>(367,475)</td>
</tr>
<tr>
<td>Intron 2 <em>g</em></td>
<td>656 G</td>
<td>E3</td>
<td>60</td>
<td>Abnormal splicing</td>
<td>SV</td>
<td>(414,418)</td>
</tr>
<tr>
<td>P105L</td>
<td>693 C→T</td>
<td>E3</td>
<td>60</td>
<td>NC</td>
<td></td>
<td>(473)</td>
</tr>
<tr>
<td>G110Δnt</td>
<td>Δ708-715</td>
<td>E3</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(485)</td>
</tr>
<tr>
<td>C168Δnt</td>
<td>991-992 TG→A</td>
<td>E4</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(367,475)</td>
</tr>
<tr>
<td>*I172N</td>
<td>1001 T→A</td>
<td>E4</td>
<td>1</td>
<td>?Insertion in ER</td>
<td>SV</td>
<td>(346,415)</td>
</tr>
<tr>
<td>G178A</td>
<td>1019 G→C</td>
<td>E5</td>
<td>0–19</td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(345)</td>
</tr>
<tr>
<td>D183E</td>
<td>1123 C→G</td>
<td>E5</td>
<td>100</td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(418)</td>
</tr>
<tr>
<td>E218E</td>
<td>Δ1160-1162</td>
<td>E5</td>
<td>6–23</td>
<td>Unstable enzyme</td>
<td></td>
<td>(437)</td>
</tr>
<tr>
<td>I236N</td>
<td>1382 T→A</td>
<td>E6</td>
<td>0</td>
<td>?Substrate binding</td>
<td>SW</td>
<td>(346,415)</td>
</tr>
<tr>
<td>V237E</td>
<td>1385 T→A</td>
<td>E6</td>
<td></td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(345)</td>
</tr>
<tr>
<td>M239K</td>
<td>1391 T→A</td>
<td>E7</td>
<td></td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(345)</td>
</tr>
<tr>
<td>S268T</td>
<td>1647 G→C</td>
<td>E7</td>
<td>100</td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(345,438)</td>
</tr>
<tr>
<td>V281L</td>
<td>1685 G→T</td>
<td>E7</td>
<td>20–50</td>
<td>?Insertion in ER</td>
<td>NC</td>
<td>(206,346)</td>
</tr>
<tr>
<td>1715 G→A</td>
<td>E7</td>
<td>0.8</td>
<td>Proton transfer from</td>
<td>H2O to heme</td>
<td>SW</td>
<td>(367,475)</td>
</tr>
<tr>
<td>1715 G→T</td>
<td>E7</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td></td>
<td>(477)</td>
</tr>
<tr>
<td>W302X</td>
<td>1750 G→A</td>
<td>E7</td>
<td></td>
<td>Abnormal splicing</td>
<td>SW</td>
<td>(478)</td>
</tr>
<tr>
<td>G110Δnt</td>
<td>1759 + T</td>
<td>E7</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(479)</td>
</tr>
<tr>
<td>Intron 7 splice donor</td>
<td>1781 G→A</td>
<td>I7</td>
<td></td>
<td>Abnormal splicing</td>
<td>SW</td>
<td>(478)</td>
</tr>
<tr>
<td>Intron 7 splice donor</td>
<td>1782 T→G</td>
<td>I7</td>
<td></td>
<td>Abnormal splicing</td>
<td>SW</td>
<td>(478)</td>
</tr>
<tr>
<td>R316X</td>
<td>1990 C→T</td>
<td>E8</td>
<td>0</td>
<td>Nonsense</td>
<td>SW</td>
<td>(416)</td>
</tr>
<tr>
<td>*Q318X</td>
<td>1996 C→T</td>
<td>E8</td>
<td>0</td>
<td>Nonsense</td>
<td>SW</td>
<td>(478)</td>
</tr>
<tr>
<td>S330Δ10nt</td>
<td>Δ2032-2041</td>
<td>E8</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(478)</td>
</tr>
<tr>
<td>R339H</td>
<td>2060 G→A</td>
<td>E8</td>
<td>20–50</td>
<td>Abnormal polymorphism</td>
<td>Normal</td>
<td>(439)</td>
</tr>
<tr>
<td>R354H</td>
<td>2105 G→A</td>
<td>E8</td>
<td>0</td>
<td>Interactions with</td>
<td>SW</td>
<td>(367,475)</td>
</tr>
<tr>
<td>R356P</td>
<td>2111 G→C</td>
<td>E8</td>
<td>0.2</td>
<td>Interactions with</td>
<td>SW</td>
<td>(366)</td>
</tr>
<tr>
<td>R356Q</td>
<td>2111 G→A</td>
<td>E8</td>
<td>1</td>
<td>Interactions with</td>
<td>SW</td>
<td>(366)</td>
</tr>
<tr>
<td>E380D</td>
<td>2267 G→T</td>
<td>E9</td>
<td>30</td>
<td>Decreased heme binding</td>
<td></td>
<td>(479,540)</td>
</tr>
<tr>
<td>V397 + 16 nt</td>
<td>Duplication</td>
<td>E9</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(478)</td>
</tr>
<tr>
<td>W405X</td>
<td>2341 G→A</td>
<td>E9</td>
<td>0</td>
<td>Nonsense</td>
<td>SW</td>
<td>(480)</td>
</tr>
<tr>
<td>G424S</td>
<td>2494 G→A</td>
<td>E10</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(481)</td>
</tr>
<tr>
<td>P453S</td>
<td>2580 C→T</td>
<td>E10</td>
<td>20–50</td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(439,472)</td>
</tr>
<tr>
<td>R475Δnt</td>
<td>Δ2649</td>
<td>E10</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(427)</td>
</tr>
<tr>
<td>R483P</td>
<td>2672 G→C</td>
<td>E10</td>
<td>1–2</td>
<td>Unstable enzyme</td>
<td></td>
<td>(482)</td>
</tr>
<tr>
<td>R483Δnt</td>
<td>2672-2723</td>
<td>E10</td>
<td>0</td>
<td>Frameshift</td>
<td>SW</td>
<td>(473)</td>
</tr>
<tr>
<td>GG→C</td>
<td></td>
<td></td>
<td></td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(345)</td>
</tr>
<tr>
<td>N493S</td>
<td>2702 A→G</td>
<td>E10</td>
<td>100</td>
<td>Normal polymorphism</td>
<td>Normal</td>
<td>(345)</td>
</tr>
</tbody>
</table>

*Δ, Deletion; +, insertion; nt, nucleotide. Single letter amino acid codes: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine; X, stop codon. As an example of mutation terminology, W22X is a nonsense mutation of tryptophan-22.

** Asterisks (*) denote mutations generated by intergenic recombinations between CYP21 and CYP21P.

* Nucleotides are numbered beginning with the A in the initial ATG of the coding sequence. Introns are included. Numbering is based on a consensus of sequences from Refs. 343–345; in general, numbers are 3 less than those in Ref. 345. A, C, G, and T are nucleotides and are not to be confused with single letter amino acid codes.

* E, Exon; I, intron.

* Activity when the mutant enzyme is expressed in cultured cells. When 2 numbers are present, the higher and lower numbers denote activity using 17-hydroxyprogesterone and progesterone as substrates, respectively. nl, Normal.

SW, Salt wasting; SV, simple virilizing; NC, nonclassic.
450–461), these deletions have been detected by genomic blot hybridization as absence (or diminished intensity in heterozygotes) of gene-specific fragments produced by digestion with several restriction enzymes (Fig. 14). Large gene conversions, in which multiple mutations are transferred when gene-specific restriction endonuclease sites are affected by gene conversion can also be difficult to distinguish from actual deletions of CYP21 in single restriction digests (e.g., Tag I).

Moreover, deletion or duplication of C4A and CYP21P (which, as mentioned, occurs frequently) can be very confusing when a C4B-CYP21 deletion or a large CYP21 gene conversion is present on the other chromosome. Thus, DNA from both parents should be examined whenever possible to confirm segregation of putative deletions or gene conversions. Reprobing the same blots with a probe for C4 is a useful measure to confirm deletions (410, 447, 452).

Deletions have also been directly documented by resolving very large restriction fragments using pulsed field gradient electrophoresis (448, 462, 463) and by high resolution fluorescent in situ hybridization (464).

In practice, genomic blot hybridizations are no longer routinely used for molecular diagnosis because they are more laborious and less informative than PCR-based techniques. However, PCR is unable to distinguish deletions from large gene conversions (see Section VI.H).

2. Nonsense and frameshift mutations. Two other mutations normally found in CYP21P completely prevent synthesis of an intact enzyme and cause salt wasting 21-hydroxylase deficiency if they occur in CYP21: the nonsense mutation in exon 7 of codon 318 (Q318X) (416) and the 8-nt deletion in exon 3 (413). The 1-nt insertion in exon 7 of CYP21P has generally not been identified as an independent mutation in patients with 21-hydroxylase deficiency.

3. A or C→G mutation in intron 2. The nucleotide 13 bp before the end of intron 2 (nt 656) is A or C in normal individuals. Mutation to G constitutes the single most frequent allele causing classic 21-hydroxylase deficiency.

This mutation causes aberrant splicing of intron 2 with retention of 19 nucleotides normally spliced out of mRNA.

### Table 5. Frequency of common mutations among 21-hydroxylase deficiency alleles in different populations

<table>
<thead>
<tr>
<th>Nationality</th>
<th>Total no. alleles</th>
<th>Del (^b)</th>
<th>Large conv (^b)</th>
<th>P30L</th>
<th>656g (^b)</th>
<th>Δ5nt (^b)</th>
<th>I172N</th>
<th>V281L</th>
<th>Q318X</th>
<th>R356W</th>
<th>Other conv, (^c)</th>
<th>Not conv, unknown (^d)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>394</td>
<td>26</td>
<td>5</td>
<td>2</td>
<td>31</td>
<td>3</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>(34,209)</td>
</tr>
<tr>
<td>Sweden</td>
<td>400</td>
<td>32</td>
<td>2</td>
<td>2</td>
<td>27</td>
<td>1</td>
<td>20</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>&lt;5</td>
<td>&lt;3</td>
<td>(420,424)</td>
</tr>
<tr>
<td>England</td>
<td>284</td>
<td>45</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>8</td>
<td></td>
<td>(431)</td>
</tr>
<tr>
<td>England</td>
<td>220</td>
<td>23</td>
<td>2</td>
<td>31</td>
<td>-</td>
<td>14</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>(425)</td>
</tr>
<tr>
<td>France</td>
<td>258</td>
<td>19</td>
<td>4</td>
<td>NA</td>
<td>21</td>
<td>3</td>
<td>9</td>
<td>17</td>
<td>4</td>
<td>NA</td>
<td>6</td>
<td>17</td>
<td>(334,419)</td>
</tr>
<tr>
<td>Finland</td>
<td>102</td>
<td>34</td>
<td>6</td>
<td>12</td>
<td>29</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td>10</td>
<td>4</td>
<td></td>
<td>(423)</td>
</tr>
<tr>
<td>Italy</td>
<td>146</td>
<td>26</td>
<td>3</td>
<td>20</td>
<td>6</td>
<td>11</td>
<td>8</td>
<td></td>
<td></td>
<td>12</td>
<td>14</td>
<td></td>
<td>(422)</td>
</tr>
<tr>
<td>Italy (south)</td>
<td>50</td>
<td>8</td>
<td>0</td>
<td>56</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
<td>(430)</td>
</tr>
<tr>
<td>Spain</td>
<td>58(^*)</td>
<td>16</td>
<td>10</td>
<td>2</td>
<td>26</td>
<td>5</td>
<td>2</td>
<td>17</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>(421)</td>
</tr>
<tr>
<td>Japan</td>
<td>102</td>
<td>18</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>3</td>
<td>26</td>
<td></td>
<td>(418)</td>
</tr>
<tr>
<td>China</td>
<td>40</td>
<td>20</td>
<td>25</td>
<td>28</td>
<td></td>
<td>8</td>
<td>10</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(426)</td>
</tr>
<tr>
<td>Chile</td>
<td>126</td>
<td>21</td>
<td>19</td>
<td>7</td>
<td></td>
<td>9</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td></td>
<td></td>
<td>(429)</td>
</tr>
<tr>
<td>Mexico</td>
<td>94</td>
<td>1</td>
<td>9</td>
<td>48</td>
<td>2</td>
<td>12</td>
<td>9</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>24</td>
<td></td>
<td>(427)</td>
</tr>
<tr>
<td>Brazil</td>
<td>74</td>
<td>8</td>
<td>14</td>
<td>25</td>
<td>1</td>
<td>19</td>
<td>4</td>
<td>11</td>
<td>8</td>
<td>5</td>
<td>15</td>
<td></td>
<td>(508)</td>
</tr>
<tr>
<td>Argentina</td>
<td>72</td>
<td>18</td>
<td>18</td>
<td>15</td>
<td></td>
<td>15</td>
<td>14</td>
<td>6</td>
<td></td>
<td></td>
<td>27</td>
<td></td>
<td>(428)</td>
</tr>
</tbody>
</table>

\(^{a}\) Studies included were large, genotyped most known mutations and included parental genotyping (thus allowing phase of different mutations to be determined). If several studies from the same group were available, the most complete and recent data were included.

\(^{b}\) See Table 4 for descriptions of mutations. Del, Deletion; Conv, conversion; 656g, “intron 2” nt 656 A/C→G; Δ5nt, G110Δ5nt. Some studies did not distinguish between deletions and large conversions, these are denoted by a single number set between these two columns.

\(^{c}\) Rarely occurring gene conversions including I236N/V237E/M239K and F306I. \(^{d}\) denotes alleles that are ambiguous due to lack of parental genotyping information; in these cases, homozygosity and hemizygosity (other chromosome carries a deletion) for a mutation cannot be distinguished.

\(^{d}\) Includes alleles carrying other mutations listed in Table 4 that do not represent gene conversions, and alleles where no mutation was identified with the methods used in that particular study.
Putative asymptomatic nt 656g homozygotes have been re-
ally, presentation of salt wasting signs is delayed until sev-
activity to permit adequate aldosterone synthesis. Occasion-
the disorder, indicating that they have insufficient enzymatic
hemizygous for this mutation have the salt wasting form of
mutation, most (but not all) patients who are homozygous or
normally spliced in the adrenal glands of patients with this
of normal enzyme might thus be synthesized.

cultured cells a small amount of normally spliced mRNA is

![FIG. 14. Top, An unequal cross-over generating a CYP21 deletion. For clarity, only the two C4 and two CYP21 genes are shown on each chromosome. When these are misaligned during meiosis, a cross-over can generate two daughter chromosomes, one with three copies of the C4-CYP21 tandem and the other with one copy. The latter is often a 21-hydroxylase deficiency allele. Bars below the genes denote representative restriction fragments that hybridize with CYP21 cDNA probes (Taq I fragments on top, Bgl II fragments below); numbers denote size in kilobases. The vertical line at one end of each bar represents the gene-specific restriction site associated with that fragment (the site at the other end of each fragment is located identically in CYP21 and CYP21P). Note that the fragments associated with the remaining gene on a chromosome with a deletion may be the same size as those normally associated with either CYP21 or CYP21P depending on the location of the gene-specific restriction sites. Bottom, Schematic illustration of results of genomic (Southern) blot hybridization from individuals carrying various rearrangements in the CYP21 genes. Digests with restriction enzymes Taq I and Bgl II are shown; fragment sizes are in kilobases. Thickness of bands is meant to correspond to fragment intensity on an autoradiogram of an actual experiment. 1, Normal; 2, homozygous for CYP21 deletion; 3, homozygous for CYP102P deletion; 4, heterozygous for CYP21P deletion; 5, heterozygous for CYP21P deletion; 6, homozygous for CYP21P duplication; 7, heterozygous for CYP21P duplication; 8, compound heterozygous for CYP21P deletions; 9, homozygous for gene conversion involving the Taq I site in CYP21 P; 10, heterozygous for gene conversion involving the Taq I site in CYP21 P. Note that lanes 2 and 9 (CYP21P deletion and gene conversion) appear identical in the Taq I but not the Bgl II digest, whereas lanes 2 and 3 (CYP21P and CYP21P deletions) can be distinguished only in the Taq I digest. (adapted from data in Ref. 410.)

resulting in a shift in the translational reading frame (414, 418). Almost all of the mRNA is aberrantly spliced, but in cultured cells a small amount of normally spliced mRNA is detected. If no other mutations were present, a small amount of normal enzyme might thus be synthesized.

Although it is not known what proportion of mRNA is normally spliced in the adrenal glands of patients with this mutation, most (but not all) patients who are homozygous or hemizygous for this mutation have the salt wasting form of the disorder, indicating that they have insufficient enzymatic activity to permit adequate aldosterone synthesis. Occasionally, presentation of salt wasting signs is delayed until several months of age in patients carrying this mutation (465). Putative asymptomatic nt 656g homozygotes have been re-
ported but in fact represent PCR typing artifacts (see Section VI.I.3)(466).

4. Pro-30→Leu (P30L). This mutation yields an enzyme with 30–60% of normal activity when expressed in cultured cells (438). However, enzymatic activity is rapidly lost when the cells are lysed, suggesting that the enzyme is relatively unstable. Patients carrying this mutation tend to have more severe signs of androgen excess than patients carrying the more common nonclassic mutation V281L (420, 438). This mutation is found in approximately one-sixth of alleles in patients with nonclassic disease, but it may comprise a higher percentage of such alleles in Japan (467).

As is the case with other microsomal P450 enzymes, CYP21 is targeted and anchored to the membrane of the endoplas-
mic reticulum mainly by a hydrophobic “tail” at the amino terminus; this tail is required for enzymatic activity and stability (468, 469). Most P450 enzymes have one or more proline residues separating this tail from the remainder of the polypeptide. These residues are predicted to create a turn in the polypeptide chain, and P30L may abolish this turn. Based on studies in other P450 enzymes, this leads to improper folding of the polypeptide and may interfere with localization in the endoplasmic reticulum (470). Indeed, the P30L mutant of CYP21 is poorly localized to the endoplasmic reticulum in some (438) but not other (435) studies.

5. Ile-172→Asn (I172N). This mutation, the only one specifically associated with the simple virilizing form of the disease, results in an enzyme with approximately 1% of normal activity (346, 365) with normal substrate affinity (Km) but reduced activity (Vmax) (346, 434). The isoleucine residue normally at this position in the E helix is strongly conserved in many different P450 enzymes, and this region of the protein in other P450s interacts with the membrane of the endo-

plasmic reticulum (471). Mutation of this hydrophobic residue to a polar residue might disrupt such an interaction, weakening the association of the enzyme with the endoplasmic reticulum, and indeed improper localization to the endoplasmic reticulum has been demonstrated in some (346) but not other (434) studies. Alternatively, the mutation may disrupt an intramolecular hydrophobic interaction stabilizing the secondary structure of the enzyme; the mutant en-
zyme is abnormally sensitive to protease digestion and doesn’t incorporate heme properly (434).

Because aldosterone is normally secreted at a rate 100-1000 times lower than that of cortisol, it is apparent that 21-
hydroxylase activity would have to decrease to very low levels before it became rate-limiting for aldosterone biosyn-
thesis. Apparently, as little as 1% of normal activity allows sufficient aldosterone synthesis to prevent significant salt wasting in most cases (see Section VI.I).

6. Ile-Val-Glu-Met-235–238→Asn-Glu-Glu-Lys (I235N/V236E/ M238K). This cluster of three missense mutations in the G helix also abolishes enzymatic activity (346, 418). Interference with substrate binding has been suggested (based on sequence conservation with cholesterol side chain cleavage enzyme, another cytochrome P450) (414) but is not supported by molecular modeling of CYP21 based on the crystal structure of CYP102.
7. Val-281→Leu (V281L). V281L occurs in all or nearly all patients with nonclassic 21-hydroxylase deficiency who carry the HLA haplotype B14:DR1, an association that is presumably due to a founder effect (see Section VI.I) (206). In certain populations (such as Jews of Eastern European origin) this is a very common genetic polymorphism with a gene frequency of more than 10% (191, 192); in contrast, direct molecular screening of normal newborns in New Zealand yielded a carrier frequency of 2% (195). Overall, approximately 70% of all nonclassic alleles carry the V281L mutation (210, 334). However, the HLA-B14:DR1-associated haplotype (and/or V281L) are less common among nonclassic CAH patients in certain ethnic groups such as Yugoslavs (193) and Japanese (467). This mutation results in an enzyme with 50% of normal activity when 17-OHP is the substrate but only 20% of normal activity for progesterone (346, 357). One study suggested that the mutant enzyme is not normally localized in the endoplasmic reticulum (346), whereas another proposed that heme binding was affected (357). As another possibility, this mutation is located in the relatively well conserved I helix, which contains residues thought to be involved in proton transfer (see above).

8. Arg-356→Trp (R356W). This mutation abolishes enzymatic activity when expressed in mammalian cells (365, 418). It is located in a region of the gene encoding the K helix of the enzyme, which suggests that the mutation affects interactions with the cytochrome P450 reductase, but this has not been demonstrated experimentally (366).

9. Other mutations. Mutations that are apparently not gene conversions (i.e., they are not usually found in CYP21P) account for 5–10% of 21-hydroxylase deficiency alleles in most populations. The most frequent of these is P453S, which occurs in a number of different populations. This suggests that CYP21P may carry P453S as an occasional polymorphism and that this mutation is transferred to CYP21 in the same way as the other mutations frequently causing 21-hydroxylase deficiency (439, 472, 473).

Novel mutations are easy to detect using automated sequencing technologies in centers with well-developed prenatal or neonatal screening programs and thus have been reported at an increased rate over the past few years. Such mutations include W22X (474), P30Q (469), G90V (367, 475), G97X (476), P105L (436, 473), G187A (367, 475), deletion of E196 (437), G291S (437, 473), G291C (367, 475), W302X (477), R316X (478), R339H (439), R354H (367, 475), R356P, R356Q (these two are independent of the R356W mutation that can be generated as a gene conversion) (366), E380D (479), W405X (480), G424S (481), and R483P (482) and frameshifts at W22 (483), Y47 (484), C168 (485), and R483 (473). Several codons including W22, P30, G291, R356, and R483 have undergone several independent mutations and thus these areas may be “hotspots” for such events. Larger rearrangements include a deletion of 10 nucleotides in exon 8 and a duplication of 16 nucleotides in exon 9 (478). Additional mutations affecting splicing include a mutation of the splice acceptor of intron 1 (474) and the splice donors of introns 2 and 7 (480).

Most of these have been reported on only one chromosome. P30Q (469), G90V and G178A (475), delE196 and G291S (437), G291C (475), R339H (439), R354H (475), R356P and R356Q (366), and R483P (437) decrease or abolish enzymatic activity, and delE196 and R483P also adversely affect enzyme stability (437); P105L acts synergistically with P453S, with which it is associated in one kindred (436).

Despite an apparently exhaustive search, mutations could be not detected in CYP21 in one patient with apparent simple virilizing 21-hydroxylase deficiency (486). The patient was homozygous for an HLA haplotype shared (on one chromosome) by a second cousin with salt wasting 21-hydroxylase deficiency who carried the 8-bp deletion in exon 3 on his other chromosome. This strongly suggests that the presumed 21-hydroxylase deficiency in the patient is genetically linked to HLA and thus to CYP21. Trivial explanations aside, this suggests that a site outside the gene may be able to significantly influence its expression (see Section VI.D.3).

10. Normal polymorphisms. Several normal polymorphisms have been detected in CYP21 in the course of initial sequencing of cloned genes by several groups (343–345, 365). An extra leucine near the N terminus (this has confused numbering of other mutations in some reports) and D183E also occur in CYP21P and presumably represent gene conversions that don’t affect activity (418). K102R (345), S268T (345, 357, 487), and N493S (345, 356) do not represent gene conversions.

G. De novo recombinations

CAH is unusual among genetic diseases in the high proportion of mutant alleles generated by intergenic recombination. Both de novo deletions (488, 489) and de novo apparent gene conversions (34, 420, 490) have been documented; the latter usually involve the intron 2 nt 656g mutation and comprise approximately 1% of 21-hydroxylase deficiency alleles. In such cases, the proband carries a mutation clearly not inherited from the genetically confirmed parents. As the frequency of 21-hydroxylase deficiency alleles in the general population is approximately 2%, the allele frequency of de novo gene conversions in intron 2 in the general population should be approximately 1 in 2 × 10^4.

De novo recombinations involving CYP21 have also been documented by PCR in sperm and leukocytes (491). Unequal crossing-over is detected only in sperm (1 in 10^5–10^6 genomes), confirming that this process takes place only during meiosis. Gene conversions, however, take place at equal frequencies in somatic cells and gametes, suggesting that gene conversions occur mainly in mitosis and that meiotic recombination (i.e., double crossing-over) contributes little, if at all, to this process. The frequency of gene conversions observed by this strategy (~1 in 10^4) is consistent with the reported rate of de novo gene conversions in patients with 21-hydroxylase deficiency.

The high rate of recombinations involving the CYP21 genes may reflect their location in the major histocompatibility complex, in which a high recombination rate between genes encoding transplantation antigens may increase the diversity of the immune response and be evolutionarily favored. The mechanism by which recombination rates might be increased is not known. It is also not known whether
deletions or gene conversions within CYP21 and CYP21P, or more generally within the 30-kb tandem duplication containing these genes, take place within certain discrete regions, or hotspots. It has been suggested that sequences resembling bacteriophage lambda chi sites, which are present at relatively high frequencies within CYP21/CYP21P, might promote recombination (415, 492). This hypothesis has not been directly tested, but a recombination between TNXA and TNXB also occurred near a chi site (446).

In addition, both CYP21 (493) and CYP21P (494) have high rates of single nucleotide polymorphisms, particularly in intron 2. The significance of this is uncertain, but it may mean that additional mechanisms other than intergenic recombination generate sequence diversity within the major histocompatibility complex.

**H. Mutation detection and approaches to prenatal diagnosis**

1. **Nonmolecular techniques.** Although it is possible to prenatally diagnose 21-hydroxylase deficiency by measuring 17-OHP levels in amniotic fluid obtained by amniocentesis, this technique cannot be used for pregnancies in which the mother takes dexamethasone to suppress the fetal adrenal (see Section V.G) unless she stops the medication for 5–7 days before the amniocentesis (309, 495).

Because 21-hydroxylase deficiency was known to be closely linked to HLA, the first alternative strategy was to HLA type the proband (i.e., a living affected child in the same family) and fetal amniocytes using serological techniques (399, 496). If these were identical, the fetus could be diagnosed with 21-hydroxylase deficiency with a high degree of confidence. This technique could not be applied in kindreds in which the proband had died or was otherwise unavailable, and new mutations, HLA homozygosity, and rare complex recombinations could confound the diagnosis (496). Moreover, serological HLA typing is a complex and expensive technique, and amniocytes must be cultured for several weeks to obtain adequate quantities (497). Nevertheless, it was used until recently in some locales (498).

Initial attempts to use the CYP21 gene itself for prenatal diagnosis relied solely on genomic blot hybridization using cDNA probes, which were able to detect only deletions and large gene conversions (i.e., ~30% of affected alleles) (305, 499–501). When it was apparent that other gene conversions accounted for most of the remaining alleles, it became feasible to carry out prenatal diagnosis by detection of a limited number of mutations. Several strategies have been used; these are considered in approximate chronological order.

2. **General considerations for molecular diagnosis.** Although it is possible to detect mutations by hybridization to genomic DNA (see below), gene amplification using PCR dramatically improves the sensitivity of these techniques. However, it was initially difficult to use PCR to detect CYP21 mutations because of the paucity of primers that would amplify CYP21 without amplifying the highly homologous CYP21P pseudogene, which already carries most of the mutations of interest. Eventually PCR conditions were identified that permitted gene-specific amplification of CYP21 in two segments (Fig. 13). For each of these segments, the CYP21-specific primer includes an 8-bp segment in exon 3 that is deleted in CYP21P (502). This strategy fails to amplify CYP21 if the gene is deleted, but deletions can be detected by conventional Southern blotting. A mutant CYP21 gene would also not be amplified if it contains a gene conversion including exon 3, but such rearrangements can be detected by a second pair of PCRs encompassing exons 1–6 and 6–10. The CYP21-specific primers for these reactions are located in exon 6, in which there is a cluster of 4 nucleotides that are mutated in CYP21P.

Conversely, “back conversion” of CYP21P to include CYP21-specific sequences could lead to spurious amplification of CYP21P and false positives. This problem is also minimized by amplification of overlapping segments.

PCR-based diagnosis may be complicated by cross-contamination of samples if rigorous controls are not implemented. Furthermore, failure to amplify one haplotype may result in misdiagnosis (466) (see Section VII.I.3). Examination of flanking microsatellite markers in all family members can minimize these problems.

Finally, it must be kept in mind that a gene conversion may be sufficiently large that it includes several mutations. If only a DNA sample from the patient is analyzed, this is impossible to distinguish from compound heterozygosity for different mutations. Therefore, both parents should also be analyzed whenever possible so as to most reliably determine the phase of different mutations (i.e., whether they lie on the same or opposite alleles). Analysis of parental alleles also permits homozygotes and hemizygotes (i.e., individuals who have a mutation on one chromosome and a deletion on the other) to be distinguished.

3. **Allele-specific oligonucleotide hybridization.** Once CYP21 cDNA had been cloned, deletions and gene conversions that affected gene-specific restriction sites could be detected by Southern blotting. As previously discussed, these accounted for approximately 25% of alleles in most populations.

Point mutations could be identified by hybridization with allele-specific oligonucleotide (ASO) probes, which are short (typically 19–21 nucleotides) single-stranded DNA segments that are centered on each polymorphic or mutant nucleotide in the gene (Fig. 15). These probes are usually radioactively labeled. Under appropriate hybridization and washing conditions, a single nucleotide mismatch is sufficient to destabilize hybridization between the probe and the gene. Each mutation could thus be tested for by duplicate hybridization with pairs of probes, each corresponding to either the normal or mutant sequence.

This technique was first applied to restriction digests of genomic DNA using either Southern blots (413, 418) or in-gel hybridization (206, 410, 415, 416). This application was at the limits of sensitivity of the technique, and it was not widely applied.

PCR makes this technique much more sensitive (34, 334, 419, 422, 428, 502–508). Amplified DNA is dotted on filters and hybridized with 18 probes corresponding to the normal and mutant sequences for each of the frequently occurring gene conversions. Because many samples can be dotted on a single filter and this process can be automated, this approach is relatively efficient when large numbers of samples are to be an-
analyzed simultaneously. However, 18 independent hybridizations are laborious for small numbers of samples as are typically encountered by laboratories performing prenatal diagnosis. This procedure is also usually performed with radioactively labeled probes, although other labeling techniques are possible. Therefore, several alternative mutation detection strategies have been used. Most require a second round of PCR after relatively long gene-specific segments have been amplified.

4. Amplification-created restriction sites. Several mutations causing 21-hydroxylase deficiency (e.g., V281L and Q318X) create or destroy restriction sites and can thus be detected in restriction digests of PCR-amplified DNA by staining agarose gels with ethidium bromide (Fig. 15). Most mutations that do not involve restriction sites can be detected by locating a PCR primer adjacent to each mutation and changing its sequence to introduce a polymorphic restriction site into the amplified segment (426, 467, 509 –513). This technique thus involves a series of second round PCRs and several different restriction digests but does not require radioactivity or specialized equipment. It has been used to screen for mutations in preimplantation embryos in a couple at risk for CAH who were undergoing in vitro fertilization (514).

5. Single-stranded conformation polymorphisms. If double-stranded DNA is denatured and then quickly returned to native conditions, it will remain in a single-stranded state with a characteristic conformation stabilized by intramolecular hydrogen

---

**Fig. 15.** Commonly employed methods for detecting mutations causing CAH. A, Restriction fragment length polymorphism (RFLP). Top, Two otherwise identical double-stranded DNA molecules differ at a single position (G-C or A-T). The former (left) is part of a GGCC sequence that is digested by a restriction endonuclease, whereas the latter (right) is part of a GGAC sequence that is not digested. Thus, fragments of different sizes are produced (center); these are detected by electrophoresis (bottom). B, Allele-specific oligonucleotide hybridization. Top, Two otherwise identical double-stranded DNA molecules differ at a single position (A-T or G-C). These are denatured and hybridized with a labeled single-stranded probe that is complementary to and will hybridize with one of these sequences; this is detected by autoradiography (bottom). C, Allele-specific PCR. Top, Two otherwise identical double-stranded DNA molecules differ at a single position (A-T or G-C). They are subjected to PCR using one primer that is complementary to both molecules (arrow at the right of each molecule), and a second primer (short strand ending in A) that is complementary only to one. Thus, only the molecule on the left can be amplified by PCR under these conditions. D, Ligase detection reaction. Top, Two otherwise identical double-stranded DNA molecules differ at a single position (A-T or G-C). They are denatured and annealed with one oligonucleotide that is complementary to both molecules and fluorescently labeled (gray line at the right of each molecule), and a second oligonucleotide (short strand ending in A) that is complementary only to one. The nick between the two oligonucleotides on the left can be ligated, but the nick between oligonucleotides on the right cannot, due to the single nucleotide mismatch. Many cycles of denaturation, annealing of oligonucleotides, and ligation can be run. The size difference between ligated and unligated molecules is detected by autoradiography.
Several different fluorescent labels can also be employed.

Typing can be performed on an automated DNA sequencer.

Oligonucleotides are fluorescently labeled, the entire geno-

Some of these disadvantages may be minimized by use of a

6. Allele-specific PCR. A single nucleotide mismatch at the

Ligation detection reaction. DNA ligase can discriminate

7. Ligation detection reaction. DNA ligase can discriminate

This type of reaction can be readily multiplexed, i.e., all 18

1. Classification of disease severity. The classification of 21-

8. Oligonucleotide arrays (“DNA chips”). Allele-specific oligo-
of prenatal diagnosis and neonatal screening, and it might serve as a useful diagnostic adjunct to ACTH stimulation tests.

The simplest way to correlate genotype and phenotype is to determine which mutations are characteristically found in each type of 21-hydroxylase deficiency. This is most informative for frequently occurring mutations. Deletions and large conversions are most often found in salt wasting patients, the intron 2 nt 656g mutation is found in both salt wasting and simple virilizing patients, I172N is characteristically seen in simple virilizing patients, and V281L and P30L are found in nonclassic patients (34, 418, 419). This distribution is consistent with the compromise in enzymatic activity conferred by each mutation (see Section VI.F).

However, patients are usually compound heterozygotes for different mutations, and so this approach has little predictive value in itself. A useful analytic strategy is to consider that 21-hydroxylase deficiency is a recessive disease, and thus the phenotype of each patient is likely to reflect his or her less severely impaired allele. If mutations are provisionally classified by the degree of enzymatic compromise—severe (also termed “type A”), moderate (type B), or mild (type C)—then one might hypothesize that salt wasting patients would have severe/severe genotypes, simple virilizing patients would have severe/moderate or moderate/moderate genotypes, and nonclassic patients would have severe/mild, moderate/mild or mild/mild genotypes. In one study of 88 families (34), these three predictions were correct in 90%, 67%, and 59% of cases, respectively. The overall correct classification rate was 79%. An expanded follow-up study of the same population (209) yielded even better results, with 177 of 197 patients (88%) being correctly classified in this manner. Similar results were obtained in other studies using the same approach (420, 527).

The salt wasting, simple virilizing, and nonclassic categories are qualitative in nature, and the distinction between simple virilizing and nonclassic disease is necessarily difficult in males in whom signs of androgen excess cannot be detected at birth. Therefore, attempts have been made to correlate genotype with quantitative measures of disease severity such as basal and ACTH-stimulated 17-OHP levels, plasma renin/urinary aldosterone ratios, and Prader genital virilization scores. In general, these are no better correlated with genotype than the broader clinical categories are. There is excellent discrimination between severe and mild genotypes, but a high degree of overlap between moderate genotypes and those either more and less affected (34).

2. Explanations for discordance of genotype and phenotype. Several explanations for the less-than-complete correspondence between genotype and phenotype are possible. The most obvious is that the severity of the disease falls on a continuum and patients with disease severity near the borders of the various classifications may easily fall on either side of these borders. Several mutations and genotypes seem to be particularly associated with this problem. First, although the intron 2 nt 656g mutation is classified as severe, it is clearly “leaky” and may yield enough normally spliced mRNA to ameliorate the enzymatic deficiency in some patients. Second, the I172N mutant has marginal enzymatic function (1% of normal), and this apparently is not always sufficient to prevent salt wasting. These two explanations accounted for 12 of 20 examples of apparent discordance between genotype and phenotype in one study (209), and significant phenotypic variation was noted in a kindred in which all five affected individuals were compound heterozygotes for these two mutations (315). Third, many patients who are discordant for genotype and phenotype are compound heterozygotes for mutations that compromise enzymatic activity to different extents (209, 420); thus it appears that some of these patients actually have in vivo enzymatic activities intermediate between those seen in patients who are homozygous for each mutation. Consistent with this idea, presumed compound heterozygotes for a classic and nonclassic allele as a group have higher stimulated 17-OHP levels than presumed homozygotes for nonclassic alleles (36).

Fourth, in studies relying on detection of known mutations, additional novel mutations within CYP21 might not be detected and might adversely affect activity.

Finally, genetic or environmental factors other than 21-hydroxylase activity may influence phenotype. As discussed in Section III.E, the degree of salt wasting tends to improve with time, even in subjects who are genetically predicted to have no 21-hydroxylase activity (84), and genetically identical siblings are occasionally discordant for severity of salt wasting (34, 83); this might reflect expression of additional 21-hydroxylase activities not encoded by CYP21 (85, 86). Similarly, genetically based variations in androgen biosynthesis or sensitivity to androgens would be expected to influence expression of signs of androgen excess (Section III.D).

3. The problem of allele dropout. Inaccurate genotyping can obviously confound genotype-phenotype correlations. An important cause of inaccurate genotyping of CYP21 is unequal PCR amplification of different alleles, sometimes termed “allele dropout.” In particular, nt 656g is sometimes preferentially amplified over the corresponding two normal alleles, 656a and especially 656c, so that heterozygous carriers of nt 656g are typed as homozygotes (466). This led to several reports of high frequencies of asymptomatic homozygotes for this mutation; such individuals were usually obligate heterozygous carriers (such as parents of patients) detected in family studies (514, 528–530). This seemed physiologically possible because this mutation activates a cryptic splice site but allows some amount of normal mRNA to be synthesized (see above). However, the carrier frequencies of nt 656g predicted by these studies were implausibly high. Moreover, the “extra” nt 656g alleles failed to segregate within kindreds (466).

Preferential amplification or allele dropout, in some instances, can be alleviated by altering the PCR amplification conditions such as lowering the KCl concentration (423, 531). In any case, confusion in critical circumstances such as prenatal diagnosis is minimized by genotyping both parents whenever possible and by concomitant use of microsatellite linkage markers (466).

J. Why is CAH so common?

Classic CAH is a relatively common inherited disease, yet it is potentially lethal if untreated. As death tends to remove...
mutant alleles from the population, the question arises as to mechanisms maintaining the carrier frequency at 1–2%. Could there be a selective advantage to heterozygosity, such as exists for sickle cell anemia and resistance to malaria? If so, one would expect to find a carrier frequency for classic CAH that exceeds that predicted from the frequency of individuals with the disease (i.e., a violation of Hardy-Weinberg equilibrium). In New Zealand, the carrier frequency estimated from direct genetic testing of 600 normal newborns—2.8%—indeed exceeds the estimate derived from the frequency of classic CAH observed in newborn screening, 1.3% (195, 223). This difference, however, falls short of statistical significance. Moreover, no mechanism for a heterozygote advantage is immediately apparent.

The observed carrier frequency might be accounted for by a series of founder effects; i.e., a mutation arising in a single ancestor in a particular population is spread through the population via one or more prolific carriers. This is the likely explanation for the high frequency of CYP21 deletions on the HLA haplotype A3;Bw47;DR7 among Northern Europeans or the V281L mutation on the HLA-B14;DR1 haplotype observed at high frequency in Ashkenazi Jews. As matings between carriers are relatively infrequent at the observed carrier frequencies, there may have been insufficient time for these mutations to be selected against, a phenomenon termed “genetic drift” (532).

New mutations may replace mutant alleles lost through death of affected individuals. As discussed in Section VI.G, the incidence of de novo gene conversions is estimated to be approximately 1 in $1 \times 10^{-4}$ (491), similar to the incidence of CAH itself, suggesting that the rates of new mutation and loss from selection may indeed balance.

A selective advantage to the carrier state might arise indirectly due to the position of CYP21 within the major histocompatibility complex. There seem to be selective advantages to heterozygosity for HLA antigens as regards disease resistance (533, 534), and this might select for heterozygosity for CYP21 mutations when these are in genetic linkage disequilibrium with particular HLA antigens. It has also been speculated (535) that gametes carrying particular HLA haplotypes might be preferentially inherited—a phenomenon termed “transmission ratio distortion”—but although this is well documented in mice (536), it has not been consistently observed in the human HLA complex.

As discussed in Section III.J, nonclassic CAH is very frequent in some populations despite putative deleterious effects on fertility, leading to analogous questions about a heterozygote advantage for this form of 21-hydroxylation deficiency (191). However, the actual prevalence of infertility in this disorder is difficult to determine due to the problem of ascertainment bias. It has been speculated that there could be a selective immunological advantage for individuals carrying the nonclassic CYP21 defect who have slightly higher cortisol levels after cosyntropin stimulation than controls. There is no evidence, however, that the response to pharmacological doses of cosyntropin has any physiological relevance to immunological function (537, 538). Obviously, the alternative explanations of founder effect, genetic drift, and a high frequency of de novo mutations all apply equally to classic and nonclassic CAH.

In summary, several explanations account for the observed frequency of mutant CYP21 alleles in the general population. Although the initial report of an increased carrier frequency of mutant CYP21 alleles (195) needs to be extended and confirmed, it is unlikely that a direct selective advantage for heterozygosity can be demonstrated.

VII. Summary

More than 90% of cases of virilizing CAH (the inherited inability to synthesize cortisol) are caused by 21-hydroxylase deficiency. Females with severe, classic 21-hydroxylase deficiency are exposed to excess androgens prenatally and are born with virilized external genitalia. Approximately three-quarters of patients cannot synthesize sufficient aldosterone to maintain sodium balance and may develop potentially fatal salt wasting crises if not treated. In the mild nonclassic form of the disorder, affected females have little or no virilization at birth, but either sex may develop signs of androgen excess postnatally. The disease is caused by mutations in the CYP21 gene encoding the steroid 21-hydroxylase enzyme. More than 90% of these mutations result from intergenic recombinations between CYP21 and the closely linked CYP21P pseudogene. Approximately 20% of these are gene deletions caused by unequal crossing over during meiosis, whereas the remainder represent transfers to CYP21 of deleterious mutations normally present in CYP21P, a process termed “gene conversion” that apparently takes place during mitosis. The degree to which each mutation compromises enzymatic activity is strongly but not completely correlated with the clinical severity of the disease in patients carrying it. Prenatal diagnosis by direct mutation detection permits prenatal treatment of affected females to minimize genital virilization. Neonatal screening by hormonal methods identifies affected children before salt wasting crises develop, reducing mortality from this condition. Glucocorticoid and mineralocorticoid replacement are the mainstays of treatment, but more rational dosing and additional therapies are being developed.

References

8. Papadopoulos V 1998 Structure and function of the peripheral-


64. Prader A 1954 Der genitalbefund beim pseudohermaphroditismus femininus der kengenitalen adrenogenitalen syndromen. Helv Paediatr Acta 9:231–248


226. Brosnan PG, Brosnan CA, Kemp SF, Domek DB, Jelly DH,


Yu AC, Grant DB 1995 Adult height in women with early-treated congenital adrenal hyperplasia (21-hydroxylase type): relation to body mass index in earlier childhood. Acta Paediatr 84:899-903


Lopes LA, Dubuis JM, Vallotton MB, Sizonenko PC 1998 Should we monitor more closely the dosage of 9a-fluorohydrocortisone in salt-losing congenital adrenal hyperplasia? J Pediatr Endocrinol Metab 11:733-737


353. Ravichandran KG, Boddupalli SS, Hasemann CA, Peterson JA,
Deisenhofer J 1993 Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450’s. Science 261:731–736


Wu DA, Chung BC 1991 Mutations of P450c21 (steroid 21-hydroxylase) at Cys428, Val281, and Ser268 result in complete, partial, or no loss of enzymatic activity, respectively. J Clin Invest 88:519–523


White PC 1987 Genetics of steroid 21-hydroxylase deficiency. Recent Prog Horm Res 43:305–336


Li Y, Lau LF 1997 Adrenocorticotropic hormone regulates the activities of the orphan nuclear receptor Nur77 through modulation of phosphorylation. Endocrinology 138:4138–4146


2872


Maurer DH, Pollack MS 1985 The use of gamma interferon to increase HLA antigen expression on cultured amniotic cells used for the prenatal diagnosis of 21-hydroxylase deficiency. Ann NY Acad Sci 458:148–155


Owerbach D, Draznin MB, Carpenter RJ, Greenberg F 1992 Pre-


