The Molecular Biology Of Hypertensive Congenital Adrenal Hyperplasia

Colin Skinner

A thesis submitted for the degree of
Doctor of Philosophy

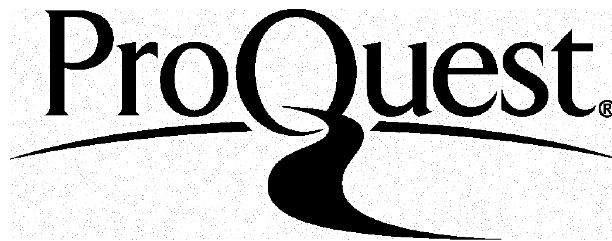
University of London

ProQuest Number: 10046198

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.



ProQuest 10046198

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The purpose of this project was to identify the molecular basis of hypertensive congenital adrenal hyperplasia in patients with two forms of the disease: steroid 17α-hydroxylase/17,20 lyase (CYP17) and 11ß-hydroxylase (CYP11B1) deficiencies. These enzyme defects account for approximately 1% and 5% of recorded cases of congenital adrenal hyperplasia respectively. CYP17 is encoded by a single gene, *CYP17* on chromosome 10. The *CYP11B1* gene is located on chromosome 8q22, in tandem with the aldosterone synthase gene, *CYP11B2*, with which it shares 93% base sequence homology. The hypothesis was proposed that the higher incidence of CYP11B1 deficiency is the result of non-homologous recombination and gene conversion between the duplicated *CYP11B* genes.

Two mutation screening methods were employed and evaluated. The first method involved construction of a genomic DNA library in bacteriophage lambda, isolation of fragments of *CYP17* and subsequent sequencing by manual dideoxy chain termination to locate mutations in DNA from a single individual with 17α-hydroxylase deficiency. The second procedure used the polymerase chain reaction (PCR) and single strand conformation polymorphism analysis (SSCP) to screen DNA from several 11ß-hydroxylase deficient patients simultaneously and proved to be a much faster and more successful approach to screening for mutations.

A single point mutation was found on each allele of the *CYP17* gene. In contrast point mutations, small deletions, small duplications and gene conversion events were found in the *CYP11B* genes, often with several mutations present in each patient. Deletion and duplication events occurred where there were direct repeats of base sequence. Pathological mutations in *CYP11B1* were shown not to arise directly from gene conversion or non-homologous recombination in the subjects studied. Other factors, such as differences in the degree of DNA methylation and selection acting strongly against mutation within *CYP17*, may account for the higher rate of mutation at the *CYP11B1* locus as compared to that of *CYP17*.

List of abbreviations

A	Adenine base
ACE	Angiotensin converting enzyme
ACTH	Adrenocorticotrophin hormone
3ß-HSD	3ß-hydroxysteroid dehydrogenase/ $\Delta^4 \rightarrow \Delta^5$ isomerase enzyme
C	Cytidine base
CAH	Congenital adrenal hyperplasia
cAMP	Cyclic adenosine monophosphate
CMO I	Corticosterone methyl oxidase I deficiency
CMO II	Corticosterone methyl oxidase II deficiency
^{m}CpG	5-methylcytosine, 5' of a guanosine nucleotide
CRE	cAMP response element
CREBP	cAMP response element binding protein
CRF	Corticotrophin releasing factor
CRS	cAMP response sequence
CYP	Cytochrome P450 enzyme
CYP	Gene encoding cytochrome P450 enzyme in all species except mouse
Cyp	Mouse gene encoding cytochrome P450 enzyme
CYP11A	Cholesterol side chain cleavage enzyme, P450scc enzyme
CYP11A	Gene encoding CYP11A enzyme
CYP11B1	Steroid 11ß-hydroxylase enzyme
CYP11B1	Gene encoding CYP11B1 enzyme
CYP11B2	Steroid 18-hydroxylase enzyme, aldosterone synthase enzyme
CYP11B2	Gene encoding CYP11B2 enzyme
CYP17	Steroid 17α-hydroxylase/ 17,20 lyase enzyme
CYP17	Gene encoding CYP17 enzyme
CYP21	Steroid 21-hydroxylase enzyme
CYP21	Gene encoding CYP21 enzyme
CYP21P	Pseudogene homologous to *CYP21*
DGGE	Denaturing gradient gel electrophoresis
DHEA	Dehydroepiandrosterone
DOC	11-Deoxycorticosterone
ER	Endoplasmic reticulum
FAD	Flavin-adenine dinucleotide
FMN	Flavin mononucleotide
G	Guanine base
GRE	Glucocorticoid response element
GRH	Glucocorticoid remediable hyperaldosteronism
HLA	Human leucocyte antigen
hsp60	60 kilodalton heat shock protein
hsp70	70 kilodalton heat shock protein

HMGCoA	3-hydroxy-3-methylglutaryl coenzyme A
17OHP	17-hydroxyprogesterone
LDL	Low density lipoproteins
MHC	Major histocompatibility complex
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
P450	Enzyme containing haem linked to protein by a cysteinyl thiolate
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pfu	Plaque forming units
PRA	Plasma renin activity
SF-1	Steroidogenic factor 1
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein
SRP	Signal recognition particle
SSCP	Single strand conformation polymorphism
T	Thymidine base
THDOC	Tetrahydrodeoxycorticosterone
THS	Tetrahydrodeoxycortisol

Table of contents

Chapter Index

Title...1

Abstract...2

List of Abbreviations...3

Table of contents - Chapter Index...5

List of Figures...9

List of Tables...13

Acknowledgements...14

Chapter 1- Introduction

1.1 Hypertension, steroid hormone production and congenital adrenal hyperplasia...15

1.2 Steroid nomenclature and molecular structure of steroid hormones...20

1.3 Adrenal steroidogenesis and cytochrome P450 enzymes

1.3.1 Cholesterol and steroidogenesis...21

1.3.2 CYP11A...23

1.3.3 CYP21...30

1.3.4 CYP17...33

1.3.5 CYP11B1 and CYP11B2...38

1.4 Diseases arising from defects in adrenal steroidogenesis

1.4.1 Congenital lipoid adrenal hyperplasia...41

1.4.2 3ß-HSD deficiency...42

1.4.3 17α-Hydroxylase/ 17,20 lyase deficiency...43

1.4.4 21-Hydroxylase deficiency...46

1.4.5 11ß-Hydroxylase deficiency...47

1.4.6 CMO I and CMO II deficiency...49

1.4.7 Glucocorticoid remediable hyperaldosteronism...50

1.5 Cytochrome P450 enzyme structure and function

1.5.1 Electron transfer and cytochrome P450 function......................51

1.5.2 Evolution by gene duplication and divergence.........................52

1.5.3 Comparison of P450 enzyme amino acid sequences and structure..56

1.6 Aims of study...58

Chapter 2- Materials and methods

2.1 Reagents

2.1.1 Stock solutions...59

2.1.2 DNA extraction solutions..61

2.1.3 Electrophoresis buffers and associated solutions.......................62

2.1.4 Southern blotting solutions..64

2.1.5 Bacterial growth media and solutions for gene library screening...66

2.1.6 Sequencing solutions...69

2.2 Methods

2.2.1 DNA extraction from blood leucocytes................................71

2.2.2 Rapid DNA extraction from 0.5 ml of whole blood................72

2.2.3 DNA preparation from buccal cells...................................73

2.2.4 Polymerase chain reaction amplification..............................74

2.2.5 Restriction endonuclease digestion.....................................77

2.2.6 Silver staining of DNA in polyacrylamide gels......................78

2.2.7 Purification of DNA from agarose....................................80

2.2.8 Southern blot analysis...83

2.2.9 Cloning DNA in bacterial vectors......................................87

2.2.10 Dideoxy chain termination DNA sequencing........................98

Chapter 3- Gene library screening for mutations in a case of 17α-hydroxylase deficiency

3.1 Clinical details...102

3.2 Experimental work

3.2.1 Southern blot analysis..102

3.2.2 Strategy for locating mutations within *CYP17*.......................103

3.2.3 *Eco*RI digestion of genomic DNA and size selection by electrophoresis..106
3.2.4 Ligation of lambda bacteriophage arms to *Eco*RI digested DNA..107
3.2.5 *In vitro* packaging of lambda bacteriophage..........................107
3.2.6 Preparation of competent C600 *hfl*A cells.............................109
3.2.7 Determination of infectivity of packaged lambda gt10.............109
3.2.8 DNA hybridization screening of a lambda gene library............110
3.2.9 Isolation of recombinant lambda gt10-*CYP17* DNA................113
3.2.10 Isolation of *CYP17* DNA and subcloning into pUC18 plasmid vector...114
3.2.11 Results of sequencing *CYP17* DNA obtained from gene library...115
3.2.12 PCR amplification of *CYP17* DNA...................................116
3.2.13 Results of sequencing PCR amplified DNA..........................116
3.3 Discussion of methodology..118
3.4 Discussion of results...119

Chapter 4- Screening for mutations in the *CYP11B* genes

4.1 Subjects with mutations in the *CYP11B* genes................................120
4.2 Strategy for locating mutations...123
4.3 SSCP analysis
4.3.1 Radioactive SSCP technique...128
4.3.2 Factors that improved SSCP resolution..............................129
4.3.3 Preparation of DNA samples for SSCP analysis....................130
4.4 Variations of SSCP conditions..133
4.5 Development of non-radioactive SSCP analysis using silver staining
4.5.1 Non-radioactive SSCP analysis..137
4.5.2 Improving silver staining SSCP technique...........................137
4.6 Results of SSCP analysis of *CYP11B1* genes
4.6.1 SSCP differences in *CYP11B1*...140
4.6.2 The *CYP11B1* gene is polymorphic....................................140
4.6.3 Unique SSCP patterns were found for all affected individuals.....142
4.7 Results of SSCP analysis of *CYP11B2* genes
4.7.1 The *CYP11B2* gene is polymorphic....................................146
4.7.2 Unique SSCP patterns from the *CYP11B2* gene.....................148

Chapter 5- Characterization of mutations in *CYP11B1* and *CYP11B2*

5.1 Sequencing *CYP11B* DNA........150

5.2 Mutations in the *CYP11B* genes

5.2.1 Small duplication in *CYP11B1* exon 2 causing frame shift........150

5.2.2 Small deletion in *CYP11B1* exon 2 causing frame shift........152

5.2.3 Nonsense mutation........152

5.2.4 Mis-sense mutations leading to amino acid changes........154

5.2.5 Intron splicing mutations........159

5.2.6 Silent exonic mutations........164

5.2.7 Gene conversions in *CYP11B1*........164

5.2.8 Gene conversions in *CYP11B2*........167

5.3 Mutations requiring verification........172

5.4 PCR artifacts and non-reproducible mutations........172

Chapter 6- Summary of results and general discussion

6.1 Mutations found in *CYP11B1*, *CYP11B2* and *CYP17* genes........173

6.2 Mechanisms of mutation

6.2.1 Slipped strand mispairing........177

6.2.2 Misalignment of repetitive DNA during homologous recombination........179

6.2.3 Deamination of methylated CpG dinucleotides in the *CYP11B* genes........181

6.2.4 Deamination of methylated CpG dinucleotides in *CYP17*........182

6.2.5 Misincorporation of nucleotides during repair or replication........183

6.2.6 Multiple mutations in *CYP11B1*........184

6.2.7 Gene conversion and non-homologous recombination........185

6.3 Mutation rate in the *CYP11B*, *CYP17* and *CYP21* genes........187

6.4 Amino acid changes in CYP11B1, CYP11B2 and CYP17........190

6.5 Evaluation of methods........196

6.6 CYP11B1, CYP17 and hypertension........200

6.7 Future directions for study of *CYP11B* and *CYP17* genes........201

6.8 Final conclusion........202

Appendix 1- PCR primers....203

Appendix 2- The three and one letter codes for amino acids....205

Appendix 3- The eukaryotic genetic code....206

Appendix 4- Published *CYP11B1* gene sequence....207

Appendix 5- Published *CYP11B2* gene sequence....210

Appendix 6- Exonic differences between *CYP11B1* and *CYP11B2* genes....214

Appendix 7- *CYP17* exonic CpG dinucleotides....217

Appendix 8- Conserved elements of P450 enzymes....219

References....227

List of Figures

Chapter 1- Introduction

1.1a The human adrenal steroidogenic pathway converting cholesterol to steroid hormones

1.1b The renin-angiotensin system controls aldosterone production, ACTH driven adrenal steroidogenesis is inhibited by cortisol

1.2a The numbering of carbon atoms and designation of ring letters for cholesterol

1.2b The molecular structures of the major adrenal steroid hormones

1.3a Transfer of hydrogen from NADPH to the flavin group

1.3b Two possible structures for the adrenodoxin [Fe_2-S_2] iron-sulphur cluster

1.3c P450 enzyme haem structure and attachment by ligation of the haem iron to a cysteinyl thiolate

1.3d The proposed mechanism of cytochrome P450 CYP11A scission of the C-20,22 bond of cholesterol

1.3e The location of *CYP21* in a duplicated region on the short arm of chromosome 6

1.3f The zonal arrangement of the adrenal cortex

1.4a Reported mutations in the *CYP17* gene

1.4b Reported mutations in the *CYP11B1* and *CYP11B2* genes

1.5a P450 enzymes act as terminal oxidases in a number of different electron transfer systems in prokaryotes and eukaryotes

1.5b The proposed P450 reaction cycle

Chapter 2- Materials and methods

2.2 Restriction map and polylinker region of the pUC18 plasmid vector

Chapter 3- Gene library screening for mutations in a case of 17α-hydroxylase deficiency

3.2a Southern blot analysis of genomic DNA following *Eco*RI digestion

3.2b Strategy for screening of a lambda gt10 gene library for *CYP17* gene fragments

3.2c Autoradiographs of primary and secondary screen blots showing positive plaques hybridizing to radioactive ^{32}P-dCTP labelled *CYP17* DNA

3.2d *Taq* cycle sequencing revealed both normal sequence, GAG, and mutant sequence, TAG, at codon 194, showing E194X mutation to be present on only one allele

3.2e Nucleotide sequences of exon 3 and 4 of *CYP17*, obtained by sequencing cloned DNA

Chapter 4- Screening for mutations in the *CYP11B* genes

4.2a Southern blot analysis following *Bam*HI digestion of genomic DNA

4.2b Strategy for screening *CYP11B* genes for mutations

4.2c Single strand conformation polymorphism (SSCP) technique for detecting single base changes within PCR amplified DNA fragments

4.3 Improvement of original SSCP technique

4.4a Effect of varying polyacrylamide concentrations on SSCP patterns

4.4b Effect of glycerol and temperature on SSCP patterns

4.5 Comparison of SSCP patterns obtained by silver staining with SSCP patterns visualized by radioactivity

4.6a-c Polymorphisms in exons 1, 3 and 4 of *CYP11B1*

4.6d SSCP of exon 2 of *CYP11B1*

4.6e SSCP of exon 4 of *CYP11B1*

4.6f SSCP of exon 6 of *CYP11B1*

4.6g SSCP of exon 8 of *CYP11B1*

4.7 SSCP variants in exon 2 and exon 9 of *CYP11B2*

Chapter 5- Characterization of mutations in *CYP11B1* and *CYP11B2*

5.2a Nucleotide sequence of *CYP11B1* exon 2 from subject 1, showing 5 bp duplication

5.2b Nucleotide sequence of *CYP11B1* exon 2 from subject 6, showing 28 bp deletion

5.2c Nucleotide sequence of *CYP11B1* exon 6 from subject 4, showing Q356X mutation

5.2d Nucleotide sequence of *CYP11B1* exon 8 from subject 1, showing R427H and V458A mutations

5.2e *Hha*I restriction digest of exon 8 DNA from normals and subject 1, confirming R427H mutation

5.2f Nucleotide sequence of *CYP11B1* exon 4 from subject 6, showing G267R mutation

5.2g Nucleotide sequence of *CYP11B1* exon 5 from subjects 7 and 8, showing G267D mutation

5.2h Nucleotide sequence of *CYP11B1* exon 9 from subjects 13 and 14, showing C494F mutation

5.2i Nucleotide sequence of *CYP11B2* exon 9 from subject 14, showing R470H mutation

5.2j Intron splicing of most nuclear mRNAs requires three consensus sequences and the binding of proteins that contain small nuclear RNAs (snRNAs)

5.2k The G267R mutation in subject 6 occurs within the left hand splice junction of intron 4 and reduces the complementarity with U1 snRNA.

5.2l Nucleotide sequence of *CYP11B1* exon 5 from subject 14 contains a G→A change that occurs at the left hand splice junction of intron 5

5.2m Nucleotide sequence of *CYP11B1* intron 1 from subjects 6 and 10, showing loss of intron lariat branch site

5.2n Nucleotide sequences of two gene conversions from *CYP11B1* to *CYP11B2*

5.2o Nucleotide sequence showing *CYP11B2* to *CYP11B1* gene conversion in intron 4 of *CYP11B2*

5.2p Gene conversion introduces a second lariat branch site for intron splicing

5.2q *Dde*I restriction digestion of *CYP11B2* intron 4/ exon 5 DNA, from the family of subjects 13 and 14, to confirm presence of gene conversion at the intron lariat branch site

5.2r *Rsa*I restriction digest of *CYP11B2* intron 4/ exon 5 DNA, from family of subjects 13 and 14 to confirm presence of gene conversion in the right hand splice consensus region

Chapter 6- Summary of results and general discussion

6.1a Mutations located in *CYP11B1*

6.1b Mutations located in *CYP17* and *CYP11B2*

6.2a Slipped strand mispairing may cause duplications in repetitive DNA

6.2b Repetitive sequences near points of deletion and duplication reported in *CYP17*, and possible causes of mutations

6.4a Proposed three dimensional structure of CYP101, a bacterial P450 from *Pseudomonas putida*, determined by X-ray crystallography

6.4b Alignment of P450 amino acid sequences; CYP101, CYP11A, CYP11B1, CYP11B2, CYP17, CYP21 and CYP102

List of Tables

Chapter 2- Materials and methods

2.2.4 DNA size markers and relative amounts of DNA within each marker band

Chapter 4- Screening for mutations in the *CYP11B* genes

4.1 Subjects with suspected mutations in *CYP11B* genes

4.2 PCR amplification fragment sizes of *CYP11B1* and *CYP11B2*, endonuclease digestion required for SSCP analysis and running times on SSCP gels

4.6.1 Numbers of different SSCP patterns obtained for each exon of *CYP11B1*

4.6.3 Exonic regions of *CYP11B1* giving a unique SSCP pattern for affected individuals

4.7.1 Numbers of different SSCP patterns obtained for each exon of *CYP11B2*

4.7.2 Distribution of unique SSCP patterns in *CYP11B2*

Chapter 5- Characterization of mutations in *CYP11B1* and *CYP11B2*

5.1 Mutations characterized in affected subjects

Acknowledgements

I would like to thank my supervisor Dr. Gillian Rumsby, for her much valued advice and patience, and my co-supervisor Dr. John Honour, for his good humour and services as general advisor on giving talks.

Great thanks also go to Cate Avey, Alison Geraerts, Richard Theaker, and Dr. Kitirat Techatraisak (and all at Chemical Pathology, Molecular Pathology and the Cobbold laboratories), both for their help and support in the lab, and for the good times both inside and outside working hours.

Acknowledgements should also go to my parents, without whom none of this would have been possible and my long-suffering wife, Monica, for the extended loan of her IBM computer, her excellent word-processing tuition, her never-ending support and the endless cups of coffee that went into making this thesis.

A final acknowledgement and heartfelt thanks should go to Dr. Peter Hindmarsh, for his initial financial support and interest in our work, and to the British Heart Foundation, whose continued funding enabled our study to be carried out.

Molecular Pathology, CAS
University College London Medical School,
Cleveland Street,
London

October 1994

1- Introduction

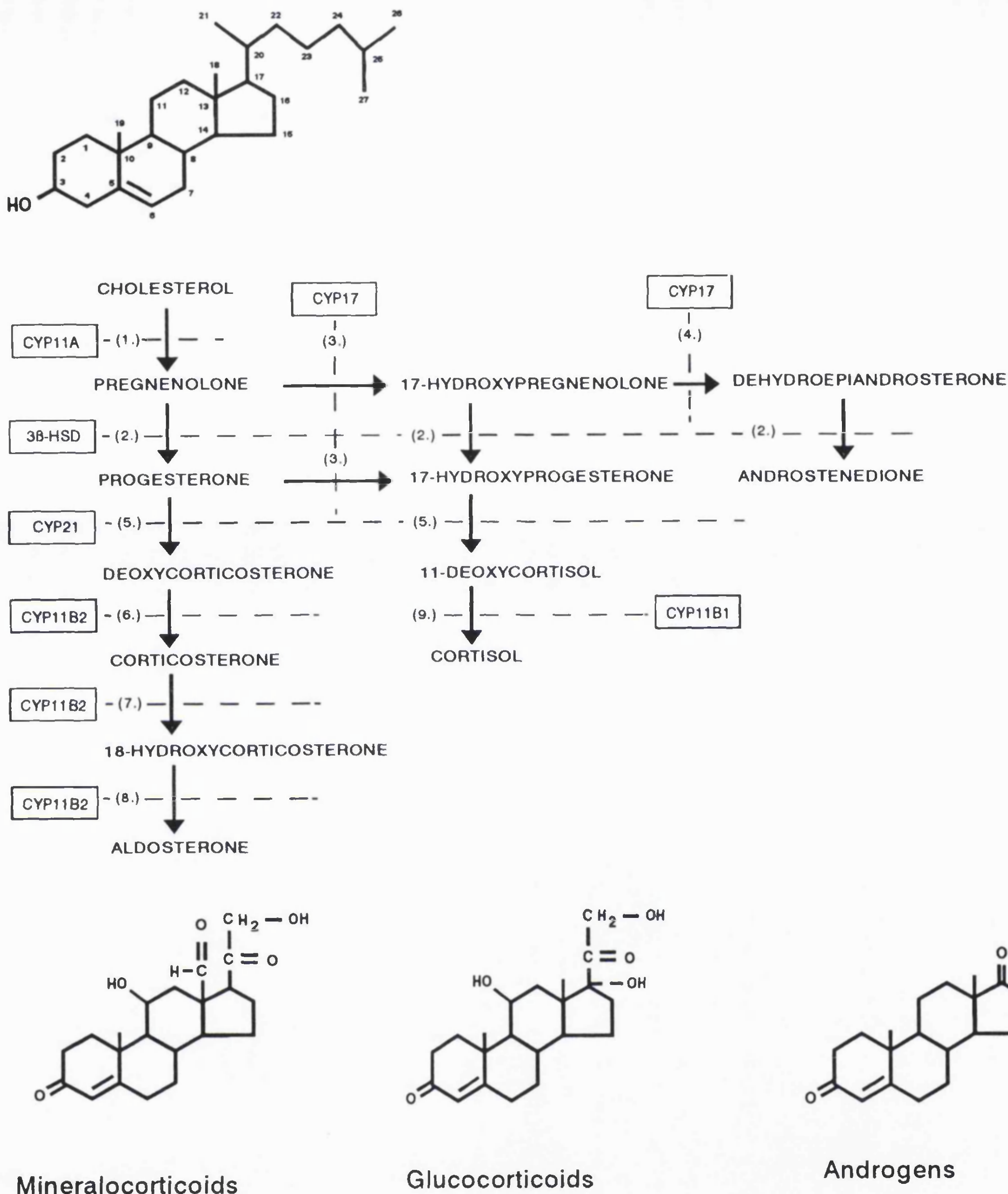
1.1 Hypertension, steroid hormone production and congenital adrenal hyperplasia

Hypertension is a complex disorder that has been linked to end-stage renal disease, brain haemorrhage and heart attack. Hypertensive disease is reported to have caused over 3,000 deaths in England and Wales in 1992* and may be a contributing factor in over 200,000 deaths per year in the United States (Hahn *et al*, 1990). Central to the cause of hypertension is the breakdown of the homeostatic mechanisms that control electrolyte balance, vaso-constriction, heart rate and circulating plasma volume. Mineralocorticoid and glucocorticoid hormones, secreted by the adrenal cortex, play a key role in the regulation of these factors. Aldosterone, the principal mineralocorticoid, regulates renal sodium retention and consequently alters electrolyte balance, intravascular volume and blood pressure. Cortisol, the principal glucocorticoid hormone, is also involved in the regulation of salt and water excretion, although its major role is in stimulating gluconeogenesis. The mineralocorticoid and glucocorticoid hormones are steroid compounds, synthesized from cholesterol by a series of enzymatic conversions (Figure 1.1a).

Two separate negative feedback mechanisms (one involving angiotensin II and the other adrenocorticotrophin hormone (ACTH)) operate to control the level of steroidogenesis. Mineralocorticoid production is regulated primarily by the renin-angiotensin system. Renin is a proteolytic enzyme which is released into the circulation when sodium levels are low or if there is a fall in extracellular fluid volume: its function is to cleave a decapeptide from angiotensinogen, an α-globulin protein present in blood plasma. This decapeptide in turn is cleaved by angiotensin converting enzyme (ACE), primarily in the lung, to produce the active octapeptide angiotensin II (O'Riordan *et al*, 1985).

* Figures supplied by Office of Population and Census and Surveys, St.Catherine's House, London.

Figure 1.1a: The human adrenal steroidogenic pathway converting cholesterol to steroid hormones.



Other quantitatively and physiologically minor steroids not shown are also produced by the same enzymes. The enzymes mediating each biosynthetic step are shown by each reaction. Reaction 1: mitochondrial CYP11A converts cholesterol to pregnenolone by catalyzing 20α-hydroxylation, 22-hydroxylation and scission of the C20-22 carbon bond. Reaction 2: 3ß-hydroxysteroid dehydrogenase (3ß-HSD) isomerase, a non-P450 enzyme bound to the endoplasmic reticulum, catalyses both 3ß-HSD and $\Delta^5 \rightarrow \Delta^4$ isomerase activities. Reactions 3 and 4: microsomal CYP17 mediates 17α-hydroxylation of pregnenolone and progesterone (3.) and the 17,20 lyase conversion of 17-hydroxypregnenolone to dehydroepiandrosterone (DHEA)(4.). Reaction 5: microsomal CYP21 catalyses 21-hydroxylation of progesterone to deoxycorticosterone (DOC) and of 17-hydroxyprogesterone to 11-deoxycortisol. Reactions 6,7,8: mitochondrial CYP11B2 catalyses three clearly distinguishable reactions: 11ß-hydroxylation (6.) of deoxycorticosterone to corticosterone, 18-hydroxylation (7.) to give 18-hydroxycorticosterone and 18-methyl oxidation (8.) to produce aldosterone. Reaction 9: mitochondrial CYP11B1, the product of a different gene than that for CYP11B2 catalyses the 11ß-hydroxylation of 11-deoxycortisol to cortisol.

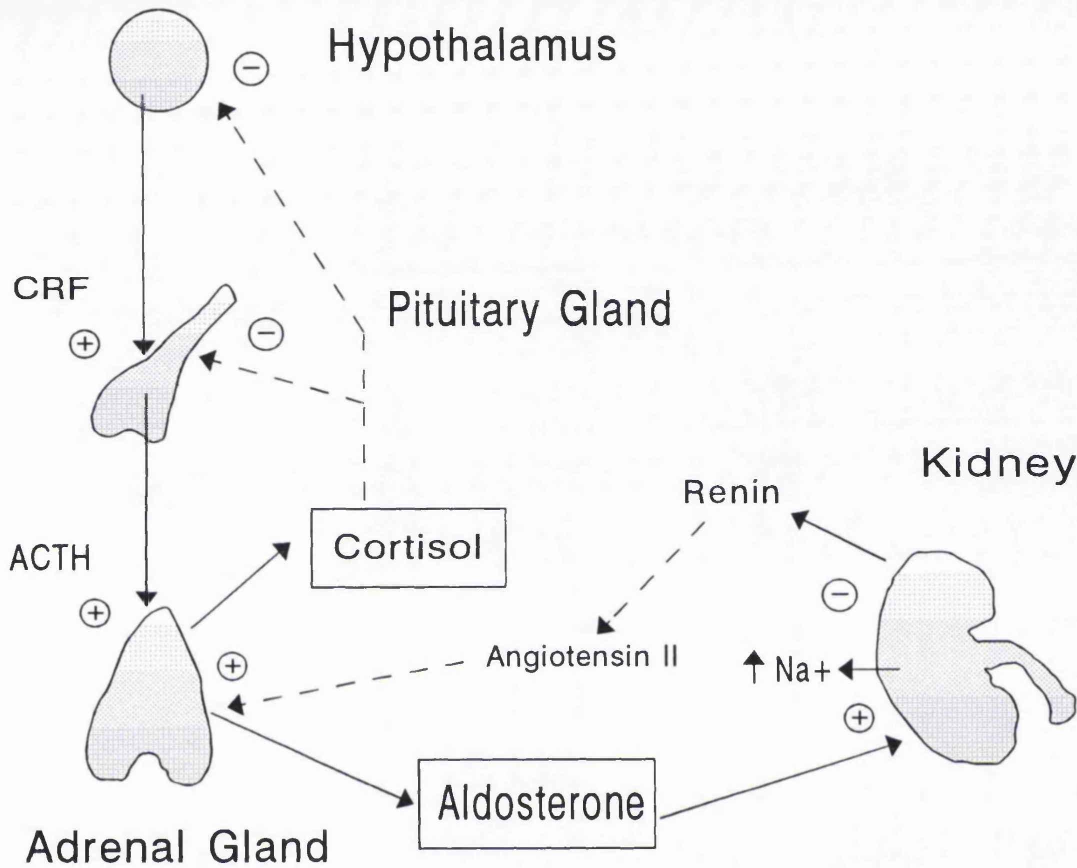
* Mineralocorticoid and glucocorticoid effects are mediated by specific cell surface receptors in different tissues. The human renal mineralocorticoid type 1 receptor mediates aldosterone action on the kidney, but these receptors are also capable of binding other steroids, such as deoxycorticosterone (DOC), 18-hydroxycorticosterone and 19-nor-deoxycorticosterone, which are not usually found at significant concentrations in human plasma, with differing affinities (Griffing *et al*, 1983). This means that renal sodium retention can be affected by several different steroid metabolites. The situation is further complicated by the fact that the renal type 1 receptors have a high affinity for cortisol. In normal subjects the microsomal enzyme 11ß-hydroxysteroid dehydrogenase (11ß-OHSD) protects the kidney from cortisol by converting the hormone to cortisone, which does not bind to mineralocorticoid receptors to any great extent. Abundant renal 11ß-OHSD carries out the conversion of the steroid to cortisone before it reaches the distal nephrons. It has been suggested that 11ß-OHSD protects various tissues from potentially deleterious effects of cortisol (Edwards *et al*, 1988; Lakshmi & Monder, 1988; Phillips *et al*, 1989). An independent enzyme, 11-oxo-reductase, acts to convert cortisone back to cortisol (Abramovitz et al, 1982; Lakshmi & Monder, 1985).

Angiotensin II binds to specific adrenocortical receptors (Sasaki *et al*, 1991) and generates a calcium signal, by release of stored calcium and increased influx of calcium through voltage dependent calcium channels (Enyedi *et al*, 1985; 1986). The increased Ca^{2+}, through the actions of calmodulin, enhances transfer of cholesterol to the inner mitochondrial membrane, activates mitochondrial dehydrogenases and stimulates the reduction of NADP, all of which contribute to increased aldosterone production. The secretion of aldosterone by adrenal glomerulosa cells is also dependent on the calcium signal (Kojima *et al*, 1984; Baukal *et al*, 1988). Angiotensin II, in addition, has a longer term effect as it induces the expression of several transcription factors (Clark *et al*, 1992) and activates protein kinase C (Kawahara *et al*, 1988; Taubman *et al*, 1989). These factors are involved in the regulation of expression of steroidogenic enzymes. The renin-angiotensin stimulation of mineralocorticoid synthesis is shut-off by negative feedback, when aldosterone binds to receptors in the kidney and causes sodium retention * that inhibits renin release (Figure 1.1b).

The production of all adrenal steroids is stimulated by ACTH. ACTH is a 39 amino acid peptide, secreted by the pituitary gland, which acts, via a receptor mediated mechanism, to mobilize free cholesterol and facilitate access of cholesterol to the side chain cleavage enzyme (CYP11A), which carries out the first, rate limiting step of steroidogenesis. ACTH, by stimulating cAMP and protein kinase A production, also activates transcription of the genes for steroidogenic enzymes and adrenocortical growth factors (John *et al*, 1986; Simpson & Waterman, 1988; Handler *et al*, 1988; Wong *et al*, 1989; Parisenti *et al*, 1993).

Inhibition of ACTH release is mediated by cortisol, which binds to specific receptors of the hypothalamus and pituitary gland. Binding of cortisol at the hypothalamus inhibits the production of corticotrophin releasing factor (CRF). As CRF acts to stimulate pituitary production of ACTH, its inhibition reduces circulating ACTH and hence steroidogenesis (Figure 1.1b). Binding of cortisol at pituitary corticotrophic cells has the same affect of inhibiting ACTH release.

Figure 1.1b: The renin-angiotensin system controls aldosterone production, and ACTH driven adrenal steroidogenesis is inhibited by cortisol.



CRF produced by the hypothalamus stimulates pituitary production and secretion of ACTH. ACTH stimulates the adrenocortical cells to transcribe genes encoding the steroidogenic enzymes and adrenal growth factors, and also stimulates cholesterol synthesis and the release of cholesterol from stored esters. The adrenal produces cortisol in response to ACTH stimulation and cortisol inhibits both hypothalamic production of CRF and pituitary production and release of ACTH, via a receptor mediated mechanism.

Renin is involved in the production of angiotensin II, which stimulates adrenal production of aldosterone. Aldosterone binds to receptors in the kidney to cause retention of sodium, plasma sodium concentration rises and renin release in inhibited. Without circulating renin, angiotensin II is cleared from the blood and aldosterone production is inhibited.

\+ denotes stimulation

\- denotes inhibition

↑ Na+ denotes the rise in plasma sodium in response to aldosterone action on the kidney.

Failure to produce cortisol results in excess circulating ACTH, excessive steroidogenesis and increased growth of the adrenal glands. In cases of congenital adrenal hyperplasia (CAH), this condition of enlarged adrenals and excess ACTH arises when individuals inherit defective genes for the enzymes involved in cortisol production. The enzyme defects prevent cortisol synthesis and the excessive ACTH drive leads to overproduction of steroids prior to the enzyme block. The clinical presentation of CAH is varied, depending upon which enzyme is deficient in the steroidogenic pathway and which precursors are overproduced.

Of particular interest in this thesis are deficiencies of two cytochrome P450 enzymes: CYP17 and CYP11B1. Deficiency of either of these enzymes can result in excessive production of the mineralocorticoid deoxycorticosterone (DOC), which leads to hypertension (Biglieri *et al*, 1966; Goldsmith *et al*, 1976; Zachmann *et al*, 1983).

The production of adrenal androgens is also affected in these two forms of CAH. 17α-hydroxylase deficiency is associated with low androgen levels, with resultant inhibition of sexual development. In contrast, 11ß-hydroxylase deficiency gives rise to excess androgen production, causing masculinization of affected females and precocious sexual development of affected males. Recent work on the molecular biology of CYP17 and CYP11B1 has identified mutations that cause loss of enzyme function (reviewed by Yanase *et al*, 1991, Curnow *et al*, 1993). Further studies have begun to elucidate the regulation of expression of these enzymes (Moore & Miller, 1991; Keeney & Waterman, 1993; Parker & Schimmer, 1993).

The gene encoding the aldosterone synthase enzyme (CYP11B2) has also been implicated in hypertension. The genes *CYP11B1* and *CYP11B2* are 93% homologous in base sequence (Mornet *et al*, 1989) and are situated in tandem on chromosome 8q21-q22 (Chua *et al*, 1987; Wagner *et al* 1991). A group of hypertensive patients has recently been reported, who possess a chimaeric *CYP11B* allele, in which the 5' end of *CYP11B1* has become joined to the 3' end of *CYP11B2* by non-homologous recombination (Pascoe *et al*, 1992a; Lifton & Dluhy, 1993). This hybrid gene encodes a functional enzyme with the steroid substrate specificity of CYP11B2 and the transcriptional regulatory elements of *CYP11B1*. A high rate of transcription of

the gene, stimulated primarily by ACTH as opposed to angiotensin II, causes excessive mineralocorticoid production, which is not inhibited by high sodium levels. Severe hypertension develops in these patients. The disorder is inherited as an autosomal dominant condition.

Although only a small proportion of hypertensive disease can be attributed to defects of CYP17, CYP11B1 and CYP11B2, the role these enzymes play in the homeostasis of salt balance and plasma volume suggests that polymorphisms and mutations within these genes, or altered levels of their expression, may predispose individuals to developing hypertension.

1.2 Steroid nomenclature and molecular structure of steroid hormones

All steroid hormones are derived from cholesterol. The carbon atoms in steroids are numbered as shown in Figure 1.2a and the four steroid rings are lettered A, B, C, and D. Cholesterol contains two methyl groups positioned at an angle to the plane of the four rings: the carbon 19 (C-19) methyl group is attached to C-10, and the C-18 methyl group is attached to C-13. A line above C-10 or C-13 denotes a methyl group. The C-18 and C-19 methyl groups of cholesterol are defined as being *above* the plane of the four rings and are said to be in a *ß-orientation* relative to this plane. A substituent that is *below* the plane of the rings is *α-oriented* and denoted by a *dashed* line. The hydroxyl group added to the steroid molecule 11-deoxycortisol by the action of the CYP11B1 enzyme is located at the C-11 position and is above the plane of the four rings, in the ß-orientation, hence the reaction is termed *11ß-hydroxylation.* In contrast, the hydroxyl group added to pregnenolone by CYP17 is located at the C-17 position and is below the plane, in the α-orientation, and so this reaction is termed *17α-hydroxylation.*

The steroid delta-5 pregnenolone (Δ^5 pregnenolone), has a double bond between C-5 and C-6 of the B ring. The action of the enzyme 3ß-hydroxysteroid dehydrogenase isomerase (3ßHSD) switches the double bond from between C-5 and C-6 in the B ring to between C-4 and C-5 in the A ring. This reaction is termed delta-5, delta-4 isomerization ($\Delta^5 \rightarrow \Delta^4$ isomerization). In addition to the isomerization reaction,

3ßHSD dehydrogenates the ß-hydroxyl group at C-3, removes a hydrogen from C-3 and adds a hydrogen to C-6. The full name of this enzyme is, therefore, 3ß-hydroxysteroid dehydrogenase/ $\Delta^5 \rightarrow \Delta^4$ isomerase, which is conveniently shortened to 3ß-HSD. The structures of the major adrenal steroid hormones and their intermediates are shown in Figure 1.2b.

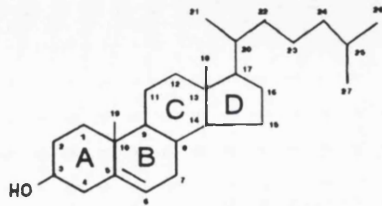
The formation of aldosterone from 11-deoxycorticosterone (DOC) requires 11ß-hydroxylation and conversion of the 18-methyl to an aldehyde group. This process is mediated by the CYP11B2 enzyme. The steroid intermediates formed in the conversion of DOC to aldosterone, corticosterone and 18-hydroxycorticosterone, are also shown in Figure 1.2b.

1.3 Adrenal steroidogenesis and cytochrome P450 enzymes

1.3.1 Cholesterol and steroidogenesis

In the process of adrenal steroid biosynthesis, cholesterol is reduced in size, hydroxylated and dehydrogenated to produce the various steroid hormones involved in the regulation of development and homeostasis. The adrenal glands can synthesize cholesterol from acetate (Brown & Goldstein, 1980), but most commonly use dietary cholesterol transported by plasma low density lipoproteins (LDLs) (Brown & Goldstein, 1986). Adequate concentrations of LDL suppress 3-hydroxy-3-methylglutaryl coenzyme A (HMGCoA) reductase, the rate limiting factor in cholesterol synthesis. The activity of HMGCoA reductase, LDL receptor synthesis and uptake of LDL are stimulated by ACTH. LDL cholesterol esters are taken up by receptor mediated endocytosis. Storage of cholesterol esters in lipid droplets is controlled by cholesterol esterase (cholesterol ester hydrolase) and cholesterol ester synthetase. ACTH stimulates the esterase and inhibits the synthetase, to increase the availability of free cholesterol for steroid hormone synthesis. The mobilization of cholesterol is thought to be the main element in the acute stimulation of steroidogenesis by ACTH (Golos & Strauss, 1988).

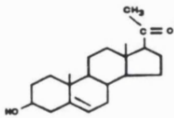
Figure 1.2a: The numbering of carbon atoms and designation of ring letters for cholesterol.



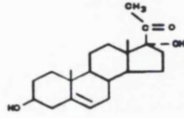
Cholesterol has 27 carbon atoms,
four rings: A, B, C and D,
a hydroxyl group at C-3,
a methyl group at C-19 and C-18,
and an eight carbon chain attached
at C-17.

Figure 1.2b: The molecular structures of the major adrenal steroid hormones.

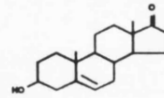
Pregnenolone



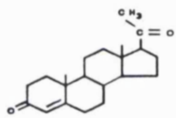
17-OH-pregnenolone



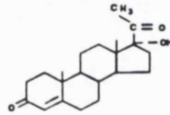
Dehydroepiandrosterone



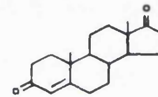
Progesterone



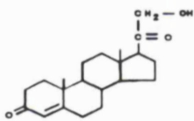
17-OH-progesterone



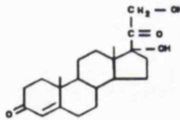
Androstenedione



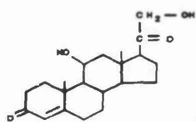
11-deoxycorticosterone



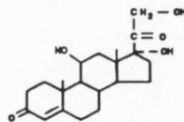
11-deoxycortisol



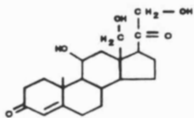
Corticosterone



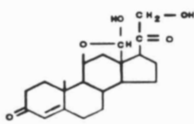
Cortisol



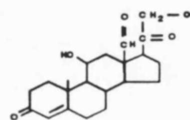
18-OH-corticosterone



Aldosterone (hemiacetal form)



Aldosterone (aldehyde form)



The transport of cholesterol into the mitochondrion appears to involve at least three proteins. Sterol carrier protein-2 (SCP-2) may bind and transport cholesterol in the cytosol and in mitochondria (Scallen *et al*, 1985). Steroidogenesis activator peptide appears to facilitate flux of cholesterol across the outer mitochondrial membrane (Pederson & Brownie, 1987), but does not seem to facilitate access of cholesterol to the cholesterol side chain cleavage enzyme, CYP11A, which acts to convert cholesterol to pregnenolone. Endozepine, the endogenous peptide ligand of the peripheral benzodiazepine receptor, appears to facilitate both transport across the outer mitochondrial membrane and access to CYP11A (Besman *et al*, 1989).

The subsequent metabolism of cholesterol then involves both mitochondrial and microsomal enzymes as described in the following sections.

1.3.2 CYP11A

Conversion of cholesterol to pregnenolone in mitochondria is the first step of steroid hormone synthesis (Figure 1.1a). The cytochrome P450 CYP11A carries out the three steps necessary for this reaction. It is encoded by a single gene, *CYP11A*, on chromosome 15q23-q24 (Chung *et al*, 1986; Morohashi *et al*, 1987; Sparkes *et al*, 1991), which is under different hormonal control in several steroidogenic tissues (Voutilainen *et al*, 1986; DiBlasio *et al*, 1987; Voutilainen & Miller, 1987; Moore & Miller, 1991).

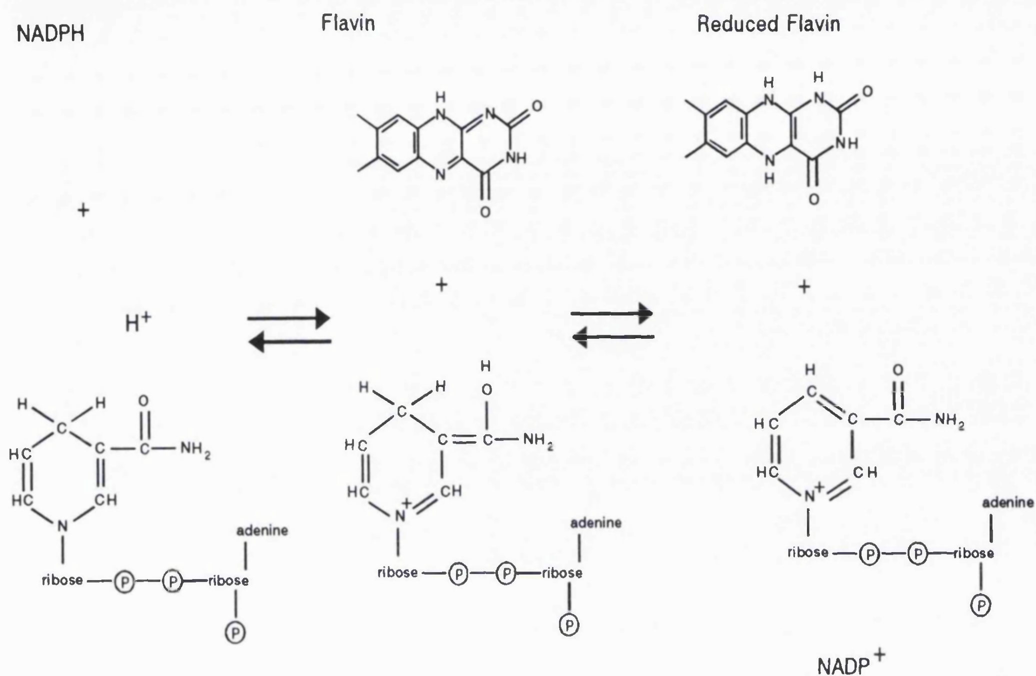
CYP11A is produced on cytosolic polyribosomes. A chaperone protein, possibly a heat-shock protein of the hsp70 family, may associate with CYP11A in the cytosol and prevent it from folding (Deshaies *et al*, 1988). An N-terminal leader peptide of around 24 amino acids is involved in targeting CYP11A to the mitochondrion. This is typical of mitochondrial matrix-targeting sequences, in that it contains several positively charged residues amongst uncharged and hydrophobic residues. Such sequences can form amphiphilic α-helices, with basic residues exposed on one face and hydrophobic residues on the other (von Heijne, 1986; Roise *et al*, 1986), which may interact with proteins on the outer mitochondrial membrane (Hartl & Neupert, 1990; Baker & Schatz, 1991).

Translocation into the mitochondria is believed to be by means of a proteinaceous pore, through which the protein passes in an extended conformation (Pfanner *et al*, 1987; Rassow *et al*, 1989). Cytosolic ATP is required, possibly to release hsp70 from the translocated protein (Hartl & Neupert, 1990). ATP and the mitochondrial proton gradient are also required to actively transport proteins into the mitochondria (Hwang & Schatz, 1989; Osterman *et al*, 1989; Rose *et al*, 1989).

Once within the mitochondrion the leader peptide is cleaved by a specific peptidase, the haem group is bound to a cysteine residue by a mitochondrial enzyme, folding of the protein and association with the inner mitochondrial membrane then proceeds (Black, 1992). A second heat shock protein of the hsp70 family, within the mitochondria, is believed to maintain translocated proteins in an unfolded conformation. ATP hydrolysis may be required to release proteins from this hsp70, so that they can associate with a 'foldase', believed to be of the hsp60 family (Kang *et al*, 1990).

CYP11A functions as the terminal oxidase in a mitochondrial electron transport system, involving NADP, the flavoprotein adrenodoxin reductase, an iron-sulphur protein, adrenodoxin, and the cytochrome P450, CYP11A (Lambeth & Pember, 1983). NADP is the initial electron acceptor in this electron transport chain. The oxidized form of NADP, $NADP^+$, acts as a coenzyme in a limited number of catabolic conversions, including the conversion of glucose-6-phosphate to ribulose-5-phosphate, and is reduced to NADPH in the process (Rose, 1979). Adrenodoxin reductase is found loosely associated with the inner mitochondrial membrane (Gnanaiah & Omdahl, 1986), and contains the flavin group FAD, which can accept hydrogen atoms from NADPH (Figure 1.3a). Adrenodoxin is a soluble protein containing an $[Fe_2\text{-}S_2]$ iron-sulphur centre and acts as an non-specific electron shuttle carrying reducing equivalents from adrenodoxin reductase to several mitochondrial P450 enzymes (Lambeth *et al*, 1979), including CYP11A, CYP11B1 and CYP11B2.

Figure 1.3a: Transfer of hydrogen from NADPH to the flavin group.



NADPH can donate two electrons to the flavin group, in a proton catalysed reaction as shown above. The flavin accepts two hydrogen atoms, which can be passed on to other electron carriers. The ability of these hydrogens to dissociate into protons and electrons facilitates electron transfer.

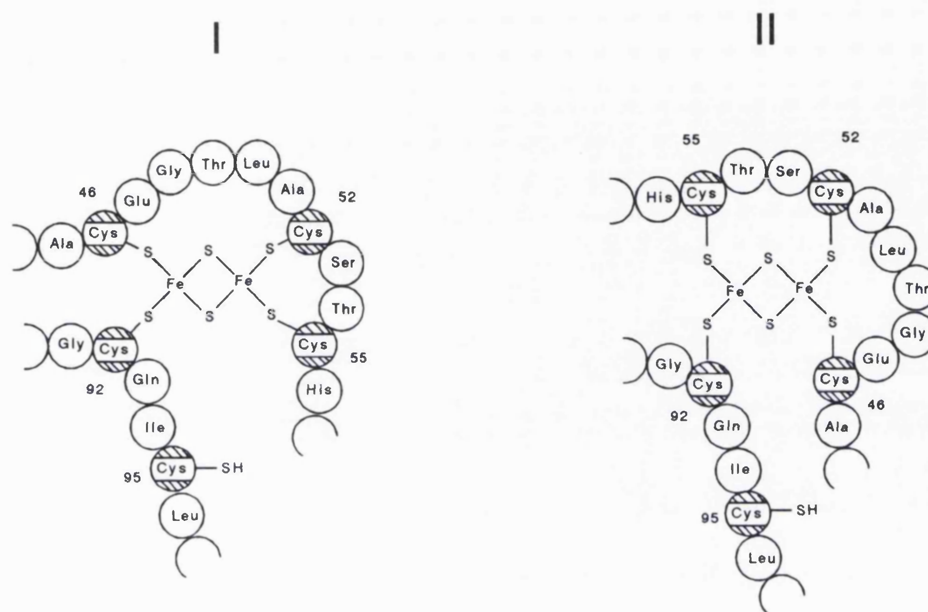
The human gene for adrenodoxin reductase is on chromosome 17q24-25, whilst the gene for adrenodoxin is on chromosome 11q22 (Sparkes *et al*, 1991). Negatively charged amino acid residues between glycine 74 and glutamic acid 86 of adrenodoxin are believed to be involved in docking with both adrenodoxin reductase and P450 enzymes (Geren *et al*, 1984; Coghlan & Vickery, 1991). The reduced flavin group of adrenodoxin reductase donates an electron to an Fe^{3+} ion in the [Fe_2-S_2] iron-sulphur cluster of adrenodoxin, by an as yet unknown mechanism. The iron-sulphur cluster is bound to adrenodoxin by four cysteines at amino acid positions 46, 52, 55 and 92 (Cupp & Vickery, 1988) in one of two possible configurations as shown in Figure 1.3b. Prior to detachment of adrenodoxin from the reductase a conformational change occurs within adrenodoxin, shifting histidine 56 (Miura & Ichikawa, 1991a; 1991b; Miura *et al*, 1991). This conformational change may lead to adrenodoxin dissociating from the reductase.

At the centre of CYP11A is an iron-containing haem group, which is ligated to a cysteinyl thiolate (Figure 1.3c), in an arrangement unique to P450 enzymes. When cholesterol is bound to a hydrophobic pocket, above the CYP11A haem, a water molecule normally ligated to the haem iron is expelled and the ligation environment of the haem alters. Induced conformational change of CYP11A is required for binding of reduced adrenodoxin to the P450 enzyme (Tsubaki *et al*, 1989). The Fe^{3+} ion of the P450 accepts an electron from adrenodoxin, to become Fe^{2+}.

The reduction of the enzyme allows molecular oxygen to complex with the haem. A second electron, and the presence of protons, is then required to split molecular oxygen and generate activated oxygen at the haem. One atom of oxygen is used to hydroxylate the steroid, while the other is reacted with two protons to form water (Coon & White, 1980). The exact mechanism is unclear, but is discussed in more detail in section 1.5.

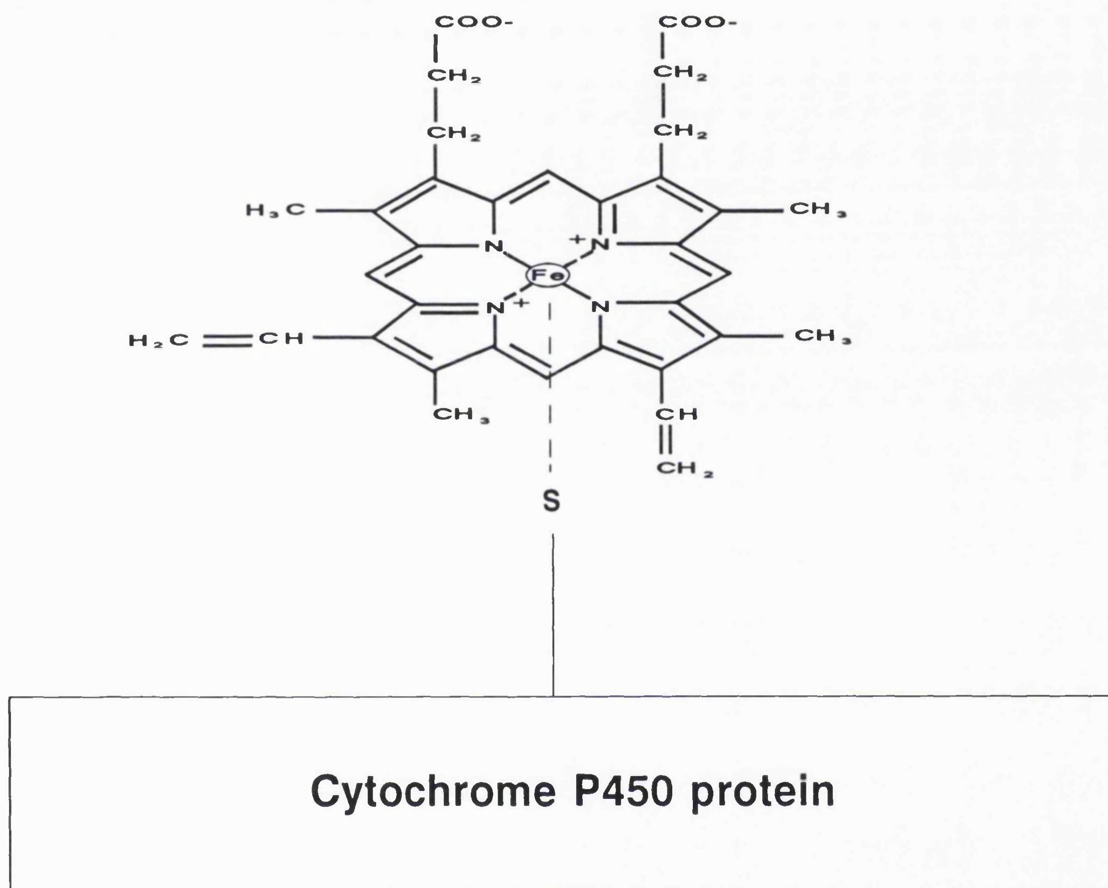
CYP11A carries out three rounds of activated oxygen formation, requiring the transfer of six electrons to the haem. The activated oxygen formed is used for 20α-hydroxylation, 22-hydroxylation and finally cleavage of the 20,22 carbon-carbon bond of cholesterol to form pregnenolone and isocaproic aldehyde (Figure 1.3d). At each step the position of the substrate above the haem is presumably altered very slightly, so that hydroxylation at two different carbon atoms and then cleavage of the bond between them can be achieved.

Figure 1.3b: Two possible structures for the adrenodoxin [Fe_2-S_2] iron-sulphur cluster (taken from Cupp & Vickery, 1988).



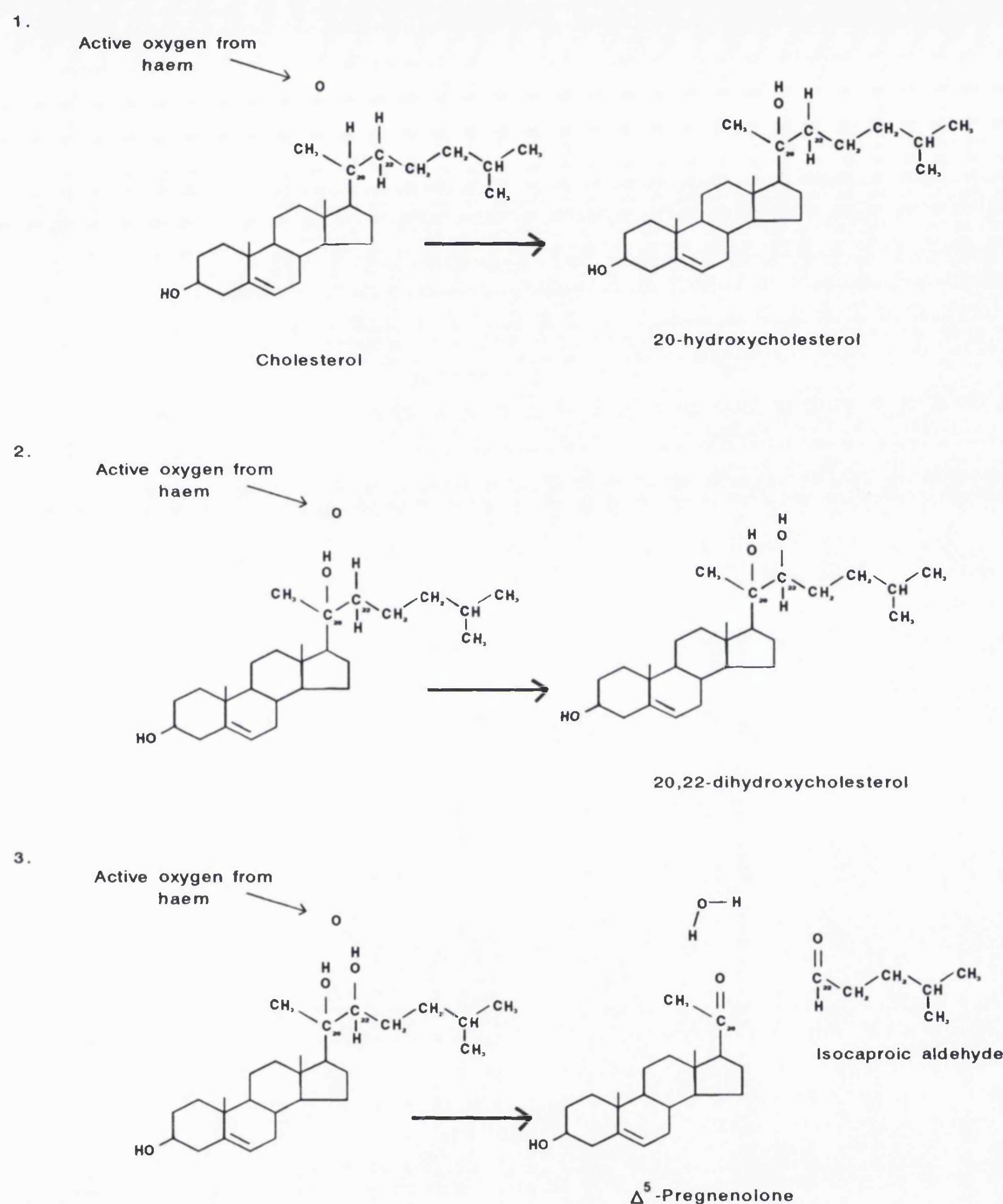
The two possible structures for the iron-sulphur centre of adrenodoxin are shown above. Structure I is based on the Cys-X_2-Cys metal binding motif, where X is a metal ion. Structure II is based on the crystal structures of ferredoxins from *S. platensis* and *Halobacterium*.

Figure 1.3c: P450 enzyme haem structure and attachment by ligation of the haem iron to a cysteinyl thiolate.



The haem group of cytochrome P450 enzymes consists of a central iron atom, which can be either an Fe^{2+} or Fe^{3+} ion, ligated to four nitrogen atoms of a porphyrin ring structure, as shown. The haem is attached to the cytochrome P450 by ligation of the haem iron to a sulphur atom of a cysteine amino acid residue of the protein. This cysteinyl thiolate is termed the fifth ligand (the four nitrogens of the porphyrin ring being the first four ligands) and a water molecule is positioned above the iron, acting as a sixth ligand, when no substrate is bound. The two propionate groups and the two vinyl groups are also involved in attachment to the cytochrome P450, by interaction with amino acid residues.

Figure 1.3d : The proposed mechanism of cytochrome P450 CYP11A scission of the C-20,22 bond of cholesterol.



1. CYP11A produces activated oxygen, O, at the haem group. The activated oxygen abstracts the hydrogen from C-20 and the hydroxyl thus formed hydroxylates C-20, to form 20α-hydroxycholesterol.

2. A similar process with a second activated oxygen atom hydroxylates C-22, to form 20,22 dihydroxycholesterol.

3. A third CYP11A mediated reaction gives rise to cleavage of the 20,22 bond. Activated oxygen abstracts the hydrogens from the 20-hydroxyl and 22-hydroxyl groups, forcing the remaining oxygens of the two hydroxyls to double bond with C-20 and C-22 and break the 20,22 bond. Water, Δ^5-pregnenolone and isocaproic aldehyde are the three resulting products.

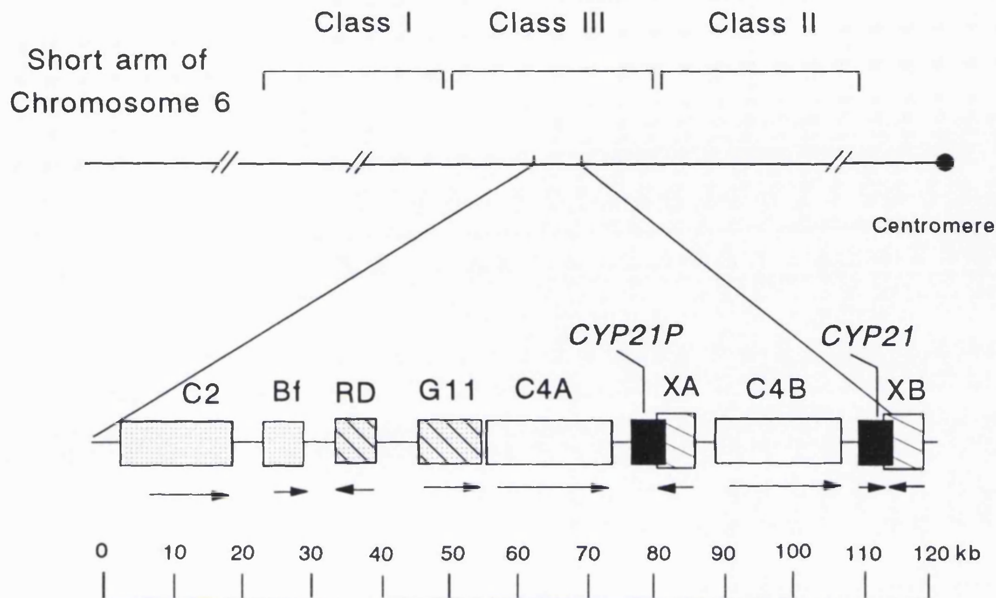
1.3.3 CYP21

Pregnenolone is converted to progesterone by an NAD dependent 3ß-hydroxysteroid dehydrogenase/ ▲5→▲4 isomerase, a non-P450 enzyme associated with the endoplasmic reticulum. Progesterone is a potent hormone involved in breast development, control of menstruation through regulation of oestrogen receptors (Hsueh *et al*, 1976), and also acts as an oestrogen antagonist. Progesterone is produced by the gonads, placenta and adrenocortical tissue. In the adrenal glomerulosa, progesterone is converted into DOC by the action of the P450 CYP21. In the zona fasciculata CYP21 acts to convert 17-hydroxyprogesterone (17OHP) into 11-deoxycortisol.

Initially it was thought that two separate 21-hydroxylase enzymes were expressed differentially in the glomerulosa and fasciculata (New *et al*, 1982). However, CYP21 has been shown to 21-hydroxylate both progesterone and 17OHP (Kominami *et al*, 1980), and cloning studies show that there is only a single functional human *CYP21* gene, which is located within the major histocompatibility (MHC) class III region of chromosome 6, in tandem with a pseudogene (*CYP21P*) (Higashi *et al*, 1986; White *et al*, 1986). The 5' region of both *CYP21P* and *CYP21* contain a duplicated C4 gene, C4A and C4B, both of which are functional and encode the fourth component of serum complement. At the 3' end of each *CYP21* gene there is an overlap with a third set of duplicated genes, XA and XB, whose coding sequence is on the opposite strand. XA and XB are believed to be structural extracellular matrix proteins (Morel *et al*, 1989). The structure of this region is shown in Figure 1.3e.

The pseudogene *CYP21P* is inactive due to an 8 bp deletion in exon 3, an insertion of a nucleotide in exon 7 and a substitution of a nucleotide in exon 8, all of which lead to premature termination codons (White *et al*, 1986). Other sequence differences between *CYP21* and the pseudogene may also prevent a functional protein being translated from *CYP21P* (see section 1.4.4).

Figure 1.3e: The location of *CYP21* in a duplicated region on the short arm of chromosome 6 (adapted from Miller, 1991).



Map of the short arm of human chromosome 6 within the region of the class I, III and II genes of the human leucocyte antigen (HLA) locus. 120 kb of the class III region is enlarged to show 10 coding regions: C2, complement factor 2; Bf, properdin factor Bf; RD, the RD 'housekeeping' gene; G11, an uncharacterized gene; C4A and C4B, nonallelic genes for complement factor 4; *CYP21P*, the pseudogene for CYP21; *CYP21*, the functional gene encoding CYP21 enzyme; XA and XB, recently discovered genes apparently encoding extracellular matrix proteins of uncertain function (Morel *et al*, 1989). The arrows beneath the genes indicate the direction of transcription, with RD, XA and XB being transcribed on the opposite strand from the other genes.

CYP21 has a hydrophobic N-terminal signal peptide of 22 amino acids, followed by lysine and arginine at positions 23 and 25 (Appendix 8). This is believed to be involved in insertion of the N-terminal of the protein into the microsomal membrane. (Wickner & Lodish, 1985). The hydrophobic signal peptide acts to anchor the protein in the microsomal membrane, while the cationic K23, R25 and possibly R16 (see Appendix 2 for amino acid 1 letter code) act as a halt-transfer sequence, which prevents the protein from passing into the lumen of the endoplasmic reticulum (Black, 1992). The translocation of ER proteins requires the signal recognition particle (SRP), SRP-receptor/ docking protein, chaperone proteins and a ribosome receptor (Wickner & Lodish, 1985; Verner & Schatz, 1988; Savitz & Meyer, 1990). These proteins appear to be involved in targeting P450 enzymes to the microsomes (Sakaguchi *et al*, 1984; Monier *et al*, 1988). The haem and active site of microsomal P450 enzymes are envisaged to be in a globular, cytoplasmic domain, rich in α-helix content. Peripheral interactions between membrane and ß-structures of P450s are also believed to take part in binding the enzymes to the microsomes (Black, 1992).

CYP21 accepts electrons from NADPH via NADPH dependent cytochrome P450 reductase (NADPH-P450 reductase), a flavoprotein containing both FAD and FMN. This is in contrast to the electron transfer system of mitochondrial P450s, which use adrenodoxin reductase and adrenodoxin as intermediate electron carriers. Human NADPH-P450 reductase is encoded by a single gene on chromosome 7 (Yamano *et al*, 1989). Electron transfer is believed to occur from NADPH→FAD→FMN→P450 (Vermilion & Coon, 1978). Hydrogen is transfered from NADPH to the flavin group of FAD, as in Figure 1.3a, then passed on to the FMN group, which contains the same flavin moiety as FAD, but not linked to adenine. The structure of reduced flavin allows FMN to donate two electrons. It is believed, however, that the FMN of NADPH P450 reductase exists in the one electron reduced form and that two molecules of reductase are required to carry out the P450 reaction (Iyanagi *et al*, 1974).

21-Hydroxylase activity is also found in a broad range of extra-adrenal tissues, especially in the fetus and during pregnancy (Casey & MacDonald, 1982; Casey *et al*, 1983), however, enzymes other than CYP21 appear to be

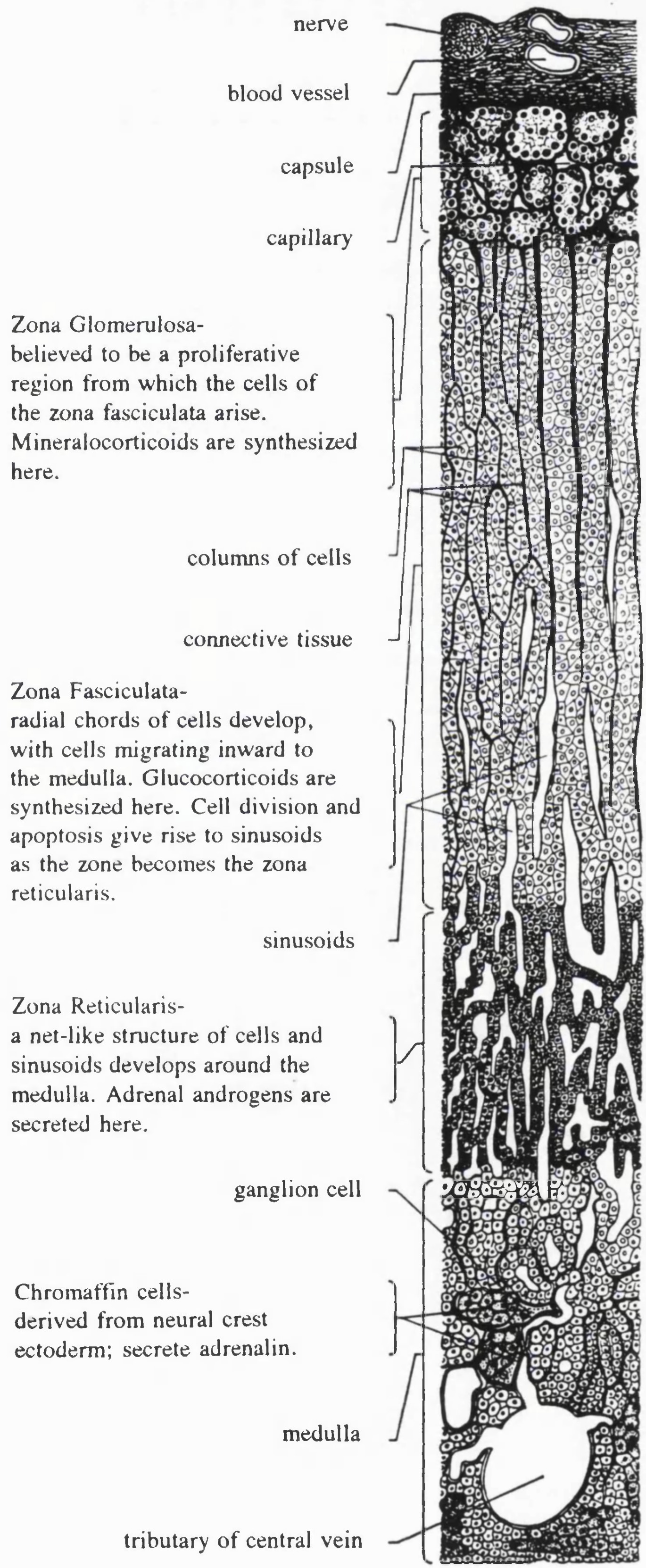
responsible (Mellon & Miller, 1989). Liver microsomal P450 enzymes, of the CYP2C family, have also been shown to carry out 21-hydroxylation in rabbits (Pendurthi *et al*, 1990). This means that complete lack of CYP21 may not lead to a total loss of 21-hydroxylated steroids (Speiser *et al*, 1991).

1.3.4 CYP17

Pregnenolone and progesterone can both be 17α-hydroxylated, to yield 17-hydroxypregnenolone and 17-hydroxyprogesterone respectively and these steroids may undergo scission of the C17,20 bond to produce dehydroepiandrosterone (DHEA) and androstenedione (Figure 1.1a). These four reactions are mediated by the same enzyme, CYP17. In the human adrenal, however, the Δ^5 pathway is favoured, so that much more 17-hydroxypregnenolone is produced. This is converted into 17OHP by the action of 3ß-HSD. Similarly the conversion of 17OHP to androstenedione does not occur, androstenedione being produced from DHEA by 3ß-HSD (Miller, 1991). In rats androstenedione is synthesized directly from 17OHP. Recent mutagenesis studies have shown that the ability of rat CYP17 to produce androstenedione from 17OHP was abolished by changing a single phenylalanine, F343, to the threonine present in human CYP17 (Koh *et al*, 1993).

CYP17 is a microsomal P450, related to CYP21 in enzyme structure and gene sequence (Picardo-Leonard & Miller, 1987). It has an N-terminal hydrophobic signal peptide that is inserted into the microsomal membrane in the same way as CYP21. CYP17 also accepts electrons from NADPH dependent cytochrome P450 reductase (Yamano *et al*, 1989). As CYP17 has both 17α-hydroxylase activity and C17,20 lyase activity it is a key branching point in steroid biosynthesis. If neither activity is present all steroids are converted through to DOC and aldosterone. This situation occurs in the outer layer of the adrenal cortex, the zona glomerulosa, where *CYP17* is not transcribed. If 17α-hydroxylation activity is present, but 17,20 lyase activity is not, as occurs in the zona fasciculata, the major steroid products are glucocorticoids. Finally, if 17α-hydroxylation and 17,20 lyase functions are both active, as occurs in the zona reticularis of the adult adrenal, androgens are produced. The transcription of *CYP17*, inhibited by angiotensin II (Mason *et al*, 1992) and stimulated by ACTH (Keeney & Waterman, 1993), and regulation of CYP17 function, are thus closely linked to the zonation of the adrenal cortex. Figure 1.3f shows the zonal arrangement of the adrenal cortex.

Figure 1.3f: The zonal arrangement of the adrenal cortex (adapted from Freeman & Bracegirdle, 1985)



The adrenals of prepubertal children synthesize abundant cortisol, but virtually no sex steroids, i.e. they have 17α-hydroxylase activity, but not 17,20 lyase activity, until adrenarche, when 17,20 lyase activity can be detected (Sklar *et al*, 1980; Mellon *et al*, 1991). This suggests separate hormonal control of the two activities. Patients have been described lacking 17,20 lyase activity, but retaining 17α-hydroxylase activity (Zachmann *et al*, 1972; Goebelsmann *et al*, 1976; Yanase *et al*, 1991). Purification of pig testicular microsomal CYP17 to homogeneity and *in vitro* reconstitution of enzyme activity showed that both 17α-hydroxylation and 17,20 lyase activities reside in a single protein (Nakajin *et al*, 1981). Cells transformed with *CYP17* cDNA acquire both activities (Zuber *et al*, 1986; Sakaki *et al* 1989; Lin *et al*, 1991a). In humans CYP17 is encoded by a single gene (*CYP17*) on chromosome 10q24-q25 (Matteson *et al*, 1986a; Chung *et al*, 1987; Sparkes *et al*, 1991).

Several factors are involved in determining whether a steroid molecule will remain on a single active site of CYP17 and undergo 17,20 bond scission following 17α-hydroxylation. CYP17 prefers Δ^5 substrates, especially for 17,20 bond scission. This is shown by the large amounts of DHEA secreted by both fetal and adult adrenal (Nakajin *et al*, 1981; Zuber *et al*, 1986; Sakaki *et al*, 1989; Lin *et al*, 1991a). Furthermore the 17α-hydroxylation occurs more readily than the 17,20 lyase reaction (Jefcoate *et al*, 1986). CYP21, CYP17 and other microsomal P450 enzymes compete for electrons from the reductase enzyme. However, perhaps more important than the relative concentrations of steroids and of competing enzymes are factors regulating electron transport to CYP17.

Increasing the ratio of NADPH dependent P450 reductase to CYP17 was shown to produce increased ratio of 17,20 lyase to 17α-hydroxylase activity (Yanagibashi & Hall, 1986). It has also been shown that cytochrome b_5 can pass electrons to CYP17. Bovine and porcine adrenal CYP17 was shown to be capable of forming a double bond at C-16, to produce 5,16-androstadien-3ß-ol (a precursor of pheromones) from pregnenolone, in the presence of testicular cytochrome b_5. The interaction was shown to be very sensitive to pH and was increased by addition of NADH- cytochrome b_5 reductase (Meadus *et al*, 1993). Cytochrome b_5 electron transfer to adrenal CYP17 has been cited as increasing 17,20 lyase activity relative to 17α-hydroxylase activity

(Onoda & Hall, 1982). This may have more effect in testis than the adrenal, however, as cytochrome b_5 is transcribed much more highly in testis (Meadus *et al*, 1993). The ability of b_5 to transfer electrons to CYP17 raises the question of whether more specific electron transfer proteins are transcribed prior to adrenarche to increase 17,20 lyase activity in the adrenals. Difference in membrane lipid association with CYP17 has also been suggested as a possible factor that could modify 17,20 lyase activity (Yanase *et al*, 1991).

The increase of 17,20 lyase function has been implicated in certain cases of functional ovarian hyperandrogenism leading to polycystic ovary syndrome (Barnes *et al*, 1993) and exaggerated adrenarche (Azziz, 1993). These conditions are distinct from hyperandrogenism due to CAH. One hypothesis for the origin of functional ovarian hyperandrogenism is that over-exposure of fetal hypothalamic-pituitary axis to maternal androgens may predispose individuals to increased *CYP17* transcription and dysregulation of 17,20 lyase activity. The inheritance of a factor affecting transcription rate of *CYP17*, or electron transfer to the enzyme could also be responsible for increased 17,20 lyase activity.

Little is known about the control of *CYP17* transcription. A putative cAMP response element, TGGAGTCA, is located at bases -131 to -124 of human *CYP17*, and sequences that resemble glucocorticoid responsive elements (GREs) may also be present (Picardo-Leonard & Miller, 1987). Bovine *CYP17* cAMP-dependent transcription is believed to be regulated by a ubiquitous protein termed cAMP response sequence 1 (CRS1), which may be specifically modified in steroidogenic cells to confer tissue specific expression (Zanger *et al*, 1992). As yet no such transcription factors have been located in humans, although there are sequence elements present in human *CYP17* that could bind similar proteins. It has been suggested that several hormonal factors and an adrenal androgen stimulating factor may act to increase the production of adrenal androgens (Parker & Odell, 1980; Parker *et al*, 1983).

Recent studies have attempted to model the CYP17 enzyme structure by alignment with CYP101 and use of computer molecular modeling (Laughton *et al*, 1990, Lin *et al*, 1993). CYP101 is one of only two bacterial proteins whose structure has been determined by X-ray crystallography (Poulos *et al*, 1985), the other being CYP102 (Ravichandran *et al*, 1993). One driving force behind the computer modeling of CYP17 is that prostatic and breast cancer have been shown to be dependent on androgens and oestrogens. The design of inhibitors which selectively knockout 17,20 lyase activity, whilst maintaining 17α-hydroxylation may be achieved by studying the active site structure of CYP17.

The low sequence homology of CYP17 and CYP101 amino acid sequences (around 15%) has meant that there is a certain degree of uncertainty about the models generated, but some interesting observations have been made. The CYP17 substrate binding pocket is described as having two lobes; one, the larger, extends from the haem in an approximately perpendicular direction, the other extends from the same point but lies almost in the plane of the haem. Both are lined mainly with hydrophobic residues, but have polar residues at the ends more remote from the haem. Both lobes are large enough to accomodate a steroid substrate such as pregnenolone (Laughton *et al*, 1990). The positioning of the two lobes suggests that the smaller of the two may be involved in 17α-hydroxylation, whilst the larger is involved in 17,20 lyase activity (Lin *et al*, 1993).

In contrast CYP21, which has structural homology with CYP17, has a single lobed substrate binding pocket. The substrate binding pockets govern which carbon atom of the steroid is reacted with oxygen at the haem.

1.3.5 CYP11B1 and CYP11B2

Mineralocorticoids and glucocorticoids are produced from DOC and 11-deoxycortisol, respectively, by the action of the cytochrome P450 11ß-hydroxylase enzymes in the mitochondria. In humans the synthesis of aldosterone is carried out by the enzyme CYP11B2, whereas cortisol is synthesized by CYP11B1 (Kawamoto *et al*, 1990; Ogishima *et al*, 1991; Curnow *et al*, 1991). These are two distinct enzymes in humans. The genes for the two enzymes (*CYP11B2* and *CYP11B1*) are arranged in tandem approximately 45 kb apart (Lifton *et al*, 1992) on chromosome 8q21-22 (Chua *et al*, 1987, Wagner *et al*, 1991). The two *CYP11B* genes are 95% homologous in their coding regions, 93% homologous in introns (Mornet *et al*, 1989), and have presumably arisen by gene duplication, similar to the *CYP21* and *CYP21P* genes. Unlike the *CYP21* genes, however, both *CYP11B1* and *CYP11B2* produce functional enzymes. There has also been divergence of the 5' regions of both *CYP11B* genes, which may reduce the chance of deletion by recombination.

Human *CYP11B2* is transcribed in the zona glomerulosa of the adrenal cortex, while *CYP11B1* is transcribed in the zona fasciculata and zona reticularis (White & Pascoe, 1992). Rats (Imai *et al*, 1990; Matsukawa *et al*, 1990) and mice (Domalik *et al*, 1991) also have distinct genes that encode isozymes active in the zonae fasciculata and glomerulosa of the adrenal cortex, whilst cattle have a single enzyme (Yanigabashi *et al*, 1986; Ohnishi *et al*, 1988) or two closely related isozymes (Ogishima *et al*, 1989; Morohashi *et al*, 1990) that are expressed in both zones and synthesize both cortisol and aldosterone, with aldosterone synthesis in the zona fasciculata of cattle being suppressed by an unknown mechanism (Wada *et al*, 1984). Rats and mice utilise corticosterone as their primary glucocorticoid. In rodents corticosterone suppression of ACTH driven steroidogenesis occurs in a similar way to that of cortisol in the human (Figure 1.1b), but may be mediated by different receptors in the hypothalamus and pituitary.

Both CYP11B1 and CYP11B2 carry an N-terminal leader peptide of around 24 amino acids, which targets them to the mitochondria, (see section 1.3.2). Both enzymes also accept electrons from NADPH via adrenodoxin reductase and adrenodoxin. In these respects the CYP11B enzymes closely resemble CYP11A. Sequence analysis suggests that all CYP11 enzymes are closely related (Mornet *et al*, 1989). Attachment to the inner mitochondrial membrane is believed to be by contacts with peripheral ß-loops of the CYP11 enzymes (Black, 1992).

Three steps are involved in the CYP11B2 catalyzed conversion of DOC to aldosterone. The first reaction adds an 11ß-hydroxyl group to DOC to form corticosterone. Corticosterone is then 18-hydroxylated. The third step requires an oxidation reaction involving the C-11 and C-18 hydroxyl groups, which forms a hemiacetal structure (Figure 1.2b). In the presence of water the hemiacetal structure can undergo a reversible reaction to give an aldehyde group at C-18. In contrast CYP11B1 is only capable of 11-hydroxylation of DOC to give corticosterone or 11-deoxycortisol to give cortisol. Both corticosterone and cortisol act as glucocorticoids, although corticosterone has a much weaker effect than cortisol in humans. CYP11B1 carries out 18-hydroxylation at about 10% the efficiency of CYP11B2, but does not produce an 18-aldehyde group to any measurable extent (Curnow *et al*, 1991; Ogishima *et al*, 1991; Kawamoto *et al*, 1992).

The divergence of transcriptional control and enzyme function is fairly recent in evolutionary terms and may provide useful insights into how duplicated genes evolve away from a common ancestral form (see Appendix 6 and section 6.2). The fact that the two proteins differ by only 41 amino acids out of a total of 503, and that many of these changes are of a conservative nature, means that key amino acids confering substrate specificity on the two enzymes should be easily located. Similarly differences in transcriptional regulation should reflect differences in control elements within the genes.

Both *CYP11B1* and *CYP11B2* contain a cAMP response element, TGACGTGA, at position -78 relative to the initiation of transcription (Mornet *et al*, 1989). A similar sequence in the somatostatin gene has been shown to bind cAMP response element binding (CREB) protein (Gonzalez & Montminy, 1989). This element may have a major effect on basal and cAMP inducible transcription of the two *CYP11B* genes, but there must be some difference to account for the fact that *CYP11B2* is expressed at very low level in response to angiotensin II, and only in the glomerulosa, while *CYP11B1* is controlled by ACTH and is highly expressed in the fasciculata (Curnow *et al*, 1991). *CYP11B1* and *CYP11B2* both contain putative binding sites for steroidogenic factor 1 (SF-1), which is a ligand inducible transcription factor found in steroidogenic tissues (Morohashi *et al*, 1992; Parker & Schimmer, 1993). The proposed SF-1 binding site of *CYP11B1* is at position -244 to -192, just 5' of a region of DNA that carries a 90 bp deletion relative to the *CYP11B2* gene and several base differences (Mornet *et al*, 1989). In *CYP11B2* there is a corresponding SF-1 site at position -271 to -324. Whether such differences within the 5' regions of the two genes can account for their differential expression is as yet unknown. Similarly, proteins that regulate *CYP11B* gene transcription have not yet been characterized.

1.4 Diseases arising from defects in adrenal steroidogenesis

1.4.1 Congenital lipoid adrenal hyperplasia

Congenital lipoid adrenal hyperplasia is the rarest and most severe form of CAH and is caused by inability to convert cholesterol to pregnenolone. Only around 30 cases have been reported, of whom only 11 survived (Hauffa *et al*, 1985). Affected males cannot produce testosterone, but produce mullerian inhibitory factor, with the result that they have female external genitalia with a blind vaginal pouch, but lack uterus, fallopian tubes and cervix. Affected females have normal genitalia at birth. Mineralocorticoid and glucocorticoid deficiencies become clinically apparent in the second week of life, with symptoms of poor feeding, lethargy, diarrhoea, vomiting, hypotension, dehydration, hyponatraemia, hyperkalaemia and acidosis. Early treatment with appropriate doses of cortisol, cortisone actetate and 9α-fluorocortisol are usually effective in correcting the disorder. Genetic males are reared as females and undergo orchiectomy in early childhood (Miller, 1991). Oestrogen replacement therapy is given at age 12 to reduce the chance of osteoporosis.

It was presumed that point mutations within the *CYP11A* gene were responsible for this condition (Matteson *et al*, 1986b), but *CYP11A* gene sequences from affected subjects analyzed to date have been normal (Lin *et al*, 1991b). The molecular basis of this disorder remains to be determined, but may be due to mutation within transcription elements or at some other locus which affects the production of pregnenolone by CYP11A enzyme.

1.4.2 3ß-HSD deficiency

There are at two main isoforms of 3ß-HSD in humans, encoded by two separate genes on chromosome 1p11-p13 (Bérubé *et al*, 1988; Lachance *et al*, 1991). Genetic analysis with exon-specific probes has revealed the presence of at least six human 3ß-HSD genes, some of which may be pseudogenes.

The two main 3ß-HSD isoforms determined to date are designated type I and type II. The type I isoform, a 372 amino acid protein encoded by a gene with four exons, is expressed primarily in the placenta and sebaceous glands of the skin. The type II isoform, 371 amino acids and highly homologous to the type I enzyme, is expressed in adrenal and gonadal tissue (Labrie *et al*, 1992).

Congenital adrenal hyperplasia due to deficiency of 3ß-HSD has been shown to be due to mutations within the type II isoform (Labrie *et al*, 1992; Rhéaume *et al*, 1992). In contrast to 21-hydroxylase and 11ß-hydroxylase deficiencies, lack of 3ß-HSD activity affects both the adrenals and gonads, resulting not only in decreased secretion of aldosterone and cortisol, but also of progesterone, androgens and oestrogens (Bongiovanni & Kellenbenz, 1962; Bongiovanni, 1981; de Peretti & Forest, 1982; New *et al*, 1989; Morel & Miller, 1991). Newborns affected by classical 3ß-HSD deficiency exhibit signs of glucocorticoid and mineralocorticoid deficiencies to various degrees, usually with severe salt-wasting that proves fatal if not diagnosed early (Pang *et al*, 1983). Affected males exhibit pseudohermaphroditism due to the impairment of androgen biosynthesis in the testis. In contrast females are are sometimes partially virilized. The female virilization occurs due to overproduction of DHEA by the foetal adrenal, which is converted to testosterone by extra-adrenal pathways involving other isoforms of 3ß-HSD (Miller, 1991). Some newborns with 3ß-HSD deficiency produce high concentrations of 17OHP, which is a product of 3ß-HSD and would expected to be produced at a lower concentration than normal. This anomaly again occurs due to the extra-adrenal production of 17-OHP by 3ß-HSD isoenzymes in skin (Cara *et al*, 1985).

Mutations found to cause classical 3ß-HSD deficiency include a homozygous nonsense mutation, of tryptophan 171 to a stop codon (W171X), and a frameshift mutation involving the insertion of a single C between codons 186 and 187 of the type II enzyme (Labrie *et al*, 1992). Both mutations occur in exon 4 of the gene and would result in a non-functional truncated protein. In all cases of 3ß-HSD, where the molecular basis of the disease has been determined, it is due to mutations in the type II isoform of the enzyme.

In addition to the classic form of 3ß-HSD there are also nonclassical, milder forms of this deficiency. Non salt-wasting forms of 3ß-HSD deficiency are assumed to be due to partial defects that allow sufficient aldosterone to be produced. Similarly the nonclassical cases may be due to mutations within the type II enzyme which affect 3ß-HSD activitiy to differing degrees.

Certain cases of hyperandrogenism may be due to mild 3ß-HSD deficiency. In these cases the overproduction of DHEA is presumed to give rise to extra-adrenal production of testosterone. In a recent study of 86 females who exhibited hirsutism and/ or hyperandrogenic oligomenorrhoea, however, deficiency of 3ß-HSD was not shown, as elevated levels of both Δ^5 and Δ^4 products were detected, which is indicative of general increase in adrenal steroidogenesis (Azziz, 1993). 3ß-HSD function in the zona reticularis is reduced compared to that in the fasciculata, whereas 17,20 lyase activity increases in the reticularis relative to the fasciculata (Labrie *et al*, 1992). This has lead to the suggestion that certain cases of female hyperandrogenism may be due to exaggerated adrenarche (Lucky *et al*, 1986).

1.4.3 17α-Hydroxylase/17,20 lyase deficiency

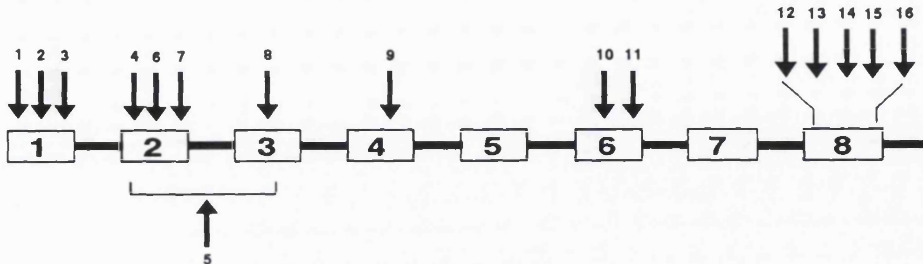
In contrast to the duplicated *CYP21* and the *CYP11B* genes, *CYP17* is a single copy gene. 17α-Hydroxylase deficiency is fairly rare, only about 120 cases having been reported (Yanase *et al*, 1991). Deficient 17α-hydroxylase activity results in decreased cortisol synthesis, with consequent overproduction of ACTH and stimulation of steps proximal to CYP17. These patients may have mild symptoms of glucocorticoid deficiency; however this is not life threatening as lack of CYP17 results in

overproduction of corticosterone, a weak glucocorticoid. These patients typically overproduce DOC in the zona fasciculata, resulting in sodium retention and hypertension. This DOC also suppresses aldosterone secretion from the zona glomerulosa so that plasma concentrations of aldosterone are not elevated, as might be inferred from Figure 1.1a. Absence of 17α-hydroxylase and 17,20 lyase activities also means that adrenal and gonadal sex steroids are not produced. As a result, genetic males may have incomplete sexual development (male pseudohermaphroditism). Affected females are phenotypically normal, but fail to undergo adrenarche or puberty. The classic presentation is that of a teenage female with sexual infantilism, amenorrhoea and hypertension (Goldsmith *et al*, 1976).

The molecular basis of 17α-hydroxylase deficiency has been determined in twenty-four patients by DNA sequencing of the mutated gene. Gross deletions of *CYP17* have not been described. The genetic lesions identified include six amino acid changes: S64Y, S106P, P342T, H373L, R440H, R496C (Imai *et al*, 1993; Lin *et al*, 1991a; Ahlgren *et al*, 1992; Monno *et al*, 1993; Fardella & Miller, 1994a; Yanase *et al*, 1991) and four premature stop codons: W17X, E194X, R239X, Q461X (Rumsby *et al*, 1992a; Ahlgren *et al*, 1992; Yanase *et al*, 1991). Deletion, duplications and insertions account for the remaining cases and comprise: a deletion of codon 53 or 54 and a 7 bp duplication at codon 120 (Yanase *et al*, 1991), duplication of codon 112 (Imai *et al*, 1993), a 4 bp duplication at codon 480 (Kagimoto *et al*, 1988), a 9 bp deletion of codons 487-489, (Fardella *et al*, 1993) and a deletion of 578 bp and insertion of 489 bp in exon 2 (Biason *et al*, 1992). These mutations are shown in (Figure 1.4a). One of these mutations, S106P, may be common in patients from Guam or other Polynesian islands (Lin *et al*, 1991a). Patients with well documented isolated 17,20 lyase deficiency have not yet been studied at the molecular level. Since they retain at least some 17α-hydroxylase activity, it is most likely that they will have point mutations causing a single amino acid change that only mildly disrupts the structure of CYP17.

Figure 1.4a: Reported mutations in the *CYP17* gene.

Mutations



	Amino acid change
1	W17X TGG->TAG
2	Deletion of TTC (Phe), codon 53 or 54
3	Y64S TAT->TCT
4	S106P TCC->CCC
6	Duplication of ATC (Ile), codon 112
8	E194X GAG->TAG
9	R239X CGA->TGA
10	P342T CCA->ACA
11	H373L CAC->CTC
12	R440H CGC->CAC
13	Q461X CAG->TAG
15	Deletion of GACTCTTTC (Asp-Ser-Phe), codons 487-489
16	R496C CGC->TGC

	Frame shift
5	Deletion of 489bp + insertion of 518bp at codon 103
7	7bp duplication of GCGCACA at codon 120
14	4bp duplication of CATC at codon 480

Mutations in the *CYP17* gene reported to cause 17α-hydroxylase deficiency are listed above. The numbered boxes represent the eight exons of *CYP17* and the intervening lines the introns.

1.4.4 21-Hydroxylase deficiency

The most common form of CAH is 21-hydroxylase deficiency, which accounts for 90-95% of CAH, and affects 1 in 12,500 live births (New *et al*, 1989). These figures are for the general population and represent children affected with the classical symptoms, certain populations have a far higher incidence of the disease and individuals with milder forms of the disease may go undetected. Perhaps 1% of the population in the Northern Hemisphere may be mildly affected heterozygotes (New *et al*, 1989). The high frequency of 21-hydroxylase deficiency is related to the structure of the *CYP21* genes of which there are two, around 30 kb apart and highly homologous, a functional gene and a pseudogene (Higashi *et al*, 1986; White *et al*, 1986). Understanding of this unique gene structure and its location in the HLA locus has lead to the definition of the molecular basis of a number of cases of 21-hydroxylase deficiency (Chiou *et al*, 1990; Higashi *et al*, 1991; Morel & Miller, 1991; Mornet *et al*, 1991). Approximately 25% of classical 21-hydroxylase deficiency cases are caused by deletion of the functional *CYP21* gene (Rumsby *et al*, 1986; Werkmeister *et al*, 1986). A proportion of these occur in association with the HLA haplotype A3 B47 DR7. This haplotype almost invariably carries a complete deletion of functional *CYP21* gene (White *et al*, 1984). Most of the remaining mutations causing 21-hydroxylase deficiency are due to point mutations, many of which are apparent gene conversion of the *CYP21* to the *CYP21P* pseudogene sequence, with loss or impairment of function of the resulting enzyme (Harada *et al*, 1987; Amor *et al*, 1988; Globerman *et al*, 1988; Higashi *et al*, 1988; Chiou *et al*, 1990).

Gene conversion is believed to occur at meiosis, when duplicated gene sequences can form heteroduplex regions of DNA. Repair enzymes recognise mismatches in the heteroduplex, excise mismatched nucleotides from one strand and replace them with nucleotides complementary to the strand from the homologous gene. The result is a non-reciprocal exchange of sequence from one gene to another, before the heteroduplex is resolved. This process can occur between homologous sequences on the same chromosome or on other chromosomes- all that is required is a short region of mismatched heteroduplex DNA (Szostak *et al*, 1983).

Due to its position in the steroid pathway CYP21 deficiency may result in loss of DOC and aldosterone production and hence salt loss, together with increased androgen synthesis. There is also a form of 21-hydroxylase deficiency where only the production of cortisol is affected, which leads to simple virilization. 21-Hydroxylase deficiency does not result in hypertension, but is of interest when looking at the generation of mutations in *CYP11B1* and *CYP11B2*, in that the two genes for the 11ß-hydroxylase isozymes are also tandemly duplicated.

1.4.5 11ß-Hydroxylase deficiency

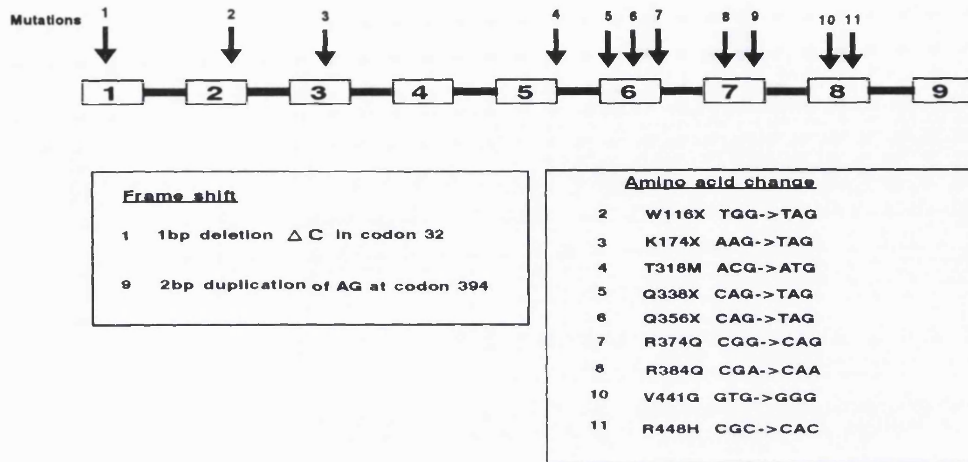
11ß-Hydroxylase deficiency occurs in about 1 in 100,000 live births, and accounts for 5-8% of congenital adrenal hyperplasia in Caucasian populations (Zachmann *et al*, 1983). In Israel, however, a large number of Jewish immigrants from Morocco are affected by the disease, with the incidence being 1 in 5,000 to 1 in 7,000 births. The majority of these cases from Israel are caused by the same mutation, a G to A change converting the arginine at position 448 of CYP11B1 to histidine (White *et al*, 1991).

As described earlier, deficiency of CYP11B1 function results in impaired production of cortisol and increased levels of ACTH, resulting in the overproduction of steroid precursors that are shunted into the pathway for androgen production. Affected females are born with masculinized external genitalia and affected individuals of both sexes undergo rapid somatic growth, but end up short due to early closure of bone epiphyses. Elevated levels of DOC cause hypertension in about two-thirds of untreated patients (Rösler *et al*, 1988).

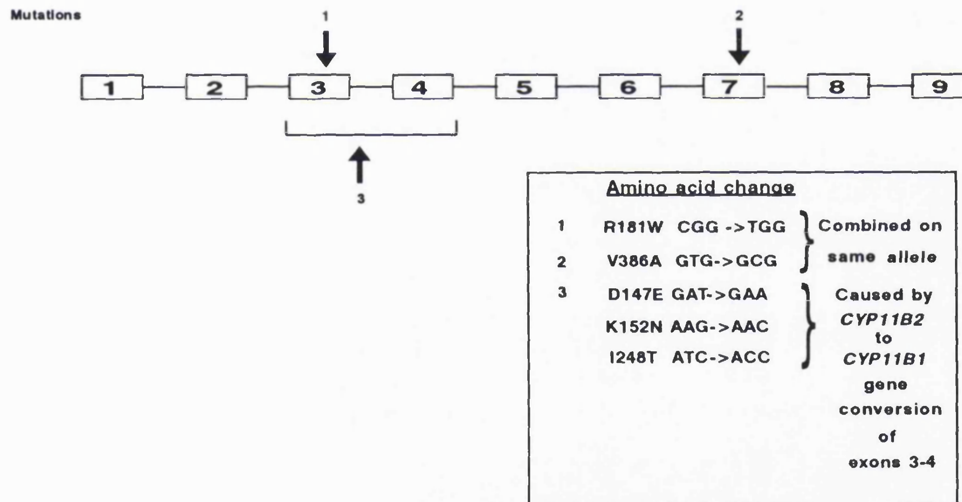
To date eleven mutations causing loss of CYP11B1 enzyme function have been described (Figure 1.4b). Of these four involve premature stop codons W116X (Naiki *et al*, 1993), K174X, Q338X, and Q356X (Curnow *et al*, 1993). Two involve frame shifts: a deletion of 1 bp from codon 32 (Curnow *et al*, 1993) and a duplication of 2 bp in codon 394 (Helmberg *et al*, 1992). The other five are amino acid changes: T318M, R374Q, R384Q, V441G (Curnow *et al*, 1993) and R448H (White *et al*, 1991).

Figure 1.4b: Reported mutations in the *CYP11B1* and *CYP11B2* genes.

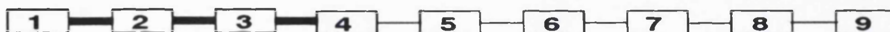
CYP11B1 (11ß-hydroxylase deficiency)



CYP11B2 (CMO II deficiency)



CYP11B1/2 (glucocorticoid remediable hyperaldosteronism)



Mutations in the *CYP11B* genes reported to cause 11ß-hydroxylase deficiency and CMO II deficiency are listed above. A third disease related to mutation within the *CYP11B* genes is termed glucocorticoid remediable hyperaldosteronism (GRH), and is caused by a fusion of the 3' end of *CYP11B2* with the 5' end of *CYP11B1*. Several cases of GRH have been characterised at the molecular level, and have been shown to be caused by non-homologous recombination at points varying from intron 2 to exon 4 of *CYP11B1*.

1.4.6 CMO I and CMO II deficiencies

Corticosterone methyloxidase (CMO) II deficiency is characterized by inability to carry out the final step in aldosterone synthesis, whereas CMO I deficiency is characterized by inability to carry out the initial 18-hydroxylation of corticosterone (Ulick *et al*, 1992). Patients with these disorders are subject to potentially fatal electrolyte balance abnormalities as neonates and a variable degree of hyponatremia and hyperkalemia combined with poor growth in childhood, but they may have no symptoms as adults (Veldhuis *et al*, 1980). CMO II deficiency is rare, except for in the population of Iranian Jews, where it was originally identified. CMO I deficiency has been reported (Ulick *et al*, 1992), but is again very rare.

In the *CYP11B2* gene two mutations have been located in patients of Iranian Jewish origin diagnosed as CMO II deficient: R181W and V386A. *CYP11B2* genes mutant for these two changes have been expressed and show that the R181W mutation causes decreased 18-hydroxylase activity and abolishes 18-oxidase activity. The V386A mutation causes a slight decrease in the function of the enzyme in converting deoxycorticosterone to aldosterone. The two mutations together cause a pronounced lack of 18-oxidase activity resulting in inability to produce aldosterone (White & Pascoe, 1992). A gene conversion altering exons 3-4 of *CYP11B2* to the sequence in *CYP11B1*, and consequently causing the amino acid changes D147E, K152N and I248T, has recently been reported to be a cause of CMO II deficiency (Fardella *et al*, 1994b). This suggests that K152 and I248 may be crucial to the *CYP11B2* ability to form aldosterone.

1.4.7 Glucocorticoid remediable hyperaldosteronism

Another set of mutations cause chimaeric genes, which contain the 5' regulatory end of *CYP11B1* and the 3' functional end of *CYP11B2* (Lifton *et al*, 1992; Pascoe *et al*, 1992a). This results in the synthesis of a protein that produces aldosterone in the zona fasciculata under the control of ACTH. In patients with these mutations hypertension occurs due to excessive production of aldosterone. The disease is termed glucocorticoid remediable hyperaldosteronism (GRH) or glucocorticoid suppressible hyperaldosteronism (GSH), as administration of the glucocorticoid dexamethasone suppresses the condition, by inhibiting the excessive ACTH drive of mineralocorticoid production. The gene for the disease acts as an autosomal dominant. In a study of 12 families affected by GRH six different cross-over break points were found in intron 2, the intron 2 exon 3 junction, exon 4 and the exon 4 intron 4 boundary. The formation of chimaeric genes seem to occur where the two sequences are most alike (Lifton *et al*, 1992).

Cyp11b2 is 5' of *Cyp11b1* in the mouse (Domalik *et al*, 1991), and studies of crossover in families with GRH suggest the same arrangement exists in humans. The reciprocal deleted chromosome in GRH patients should carry a single *CYP11B1* gene with a 5' end corresponding to *CYP11B2* and a 3' end corresponding to *CYP11B1*. Like *CYP11B2*, such a gene should be expressed at low levels and only in the zona glomerulosa, but it should have an enzyme activity similar to the normal product of *CYP11B1*. As yet, such chromosomes have not been detected (White & Pascoe, 1992).

1.5 Cytochrome P450 enzyme structure and function

1.5.1 Electron transfer and cytochrome P450 function

About thirty-five years ago hepatic microsomes were found to contain a previously unrecognized carbon-monoxide binding pigment (Ryan & Engel, 1957; Klingenberg, 1958; Garfinkel, 1958). The inhibition of oxygen dependent microsomal drug and steroid hydroxylation by carbon-monoxide had been shown to be reversible by absorption of light at 450 nm (Estabrook *et al*, 1963). This was the same wavelength at which the isolated hepatic microsomal pigment absorbed light, when reduced in the presence of carbon-monoxide. Omura and Sato (1964) named the pigment P450 and determined that it was a cytochrome of the *b* type, with a haem group that was not covalently bound to the protein. This haem was reducible by NADPH under anaerobic conditions and rapidly reoxidizable in the presence of molecular oxygen.

Certain P450 enzymes were found to transfer electrons to cytochrome b_5 (Peterson *et al*, 1977) or to substrates that serve as terminal electron acceptors, such as nitro or azo compounds (Gillette, 1966), epoxides (Kato *et al*, 1976a) and amine oxides (Kato *et al*, 1976b). In general, however, the P450 enzymes acted as oxygenases, using donated electrons to split molecular oxygen and hydroxylate or oxidize hydrophobic substrates.

Early studies on hepatic microsome P450 enzymes revealed that electrons were accepted, one at a time, from an ancillary protein, NADPH dependent cytochrome P450 reductase, which contained FAD and FMN, with electrons passing from NADPH→FAD→FMN→P450 (Vermilion & Coon, 1978). Bacterial and mitochondrial P450 enzymes were shown to accept electrons from NADH or NADPH via two intermediate proteins: an FAD containing flavoprotein and an iron-sulphur protein; electrons passing from NADH or NADPH→FAD→Fe-S→P450 haem (Gunsalus *et al*, 1973; Lambeth *et al*, 1979, Lambeth & Pember, 1983). This divides P450s into two clear groups depending on whether they accept electrons from NADPH-dependent P450 reductase or an iron-sulphur protein. The division separated eukaryotic microsomal P450s, such as CYP17 and CYP21, from the mitochondrial P450s,

including the three CYP11 enzymes, the genes for which are presumed to have originated in the prokaryotic organisms that gave rise to mitochondria, before being sequestered into the eukaryote genome. A number of variations within the two basic P450 classes have since been characterized as shown in Figure 1.5a.

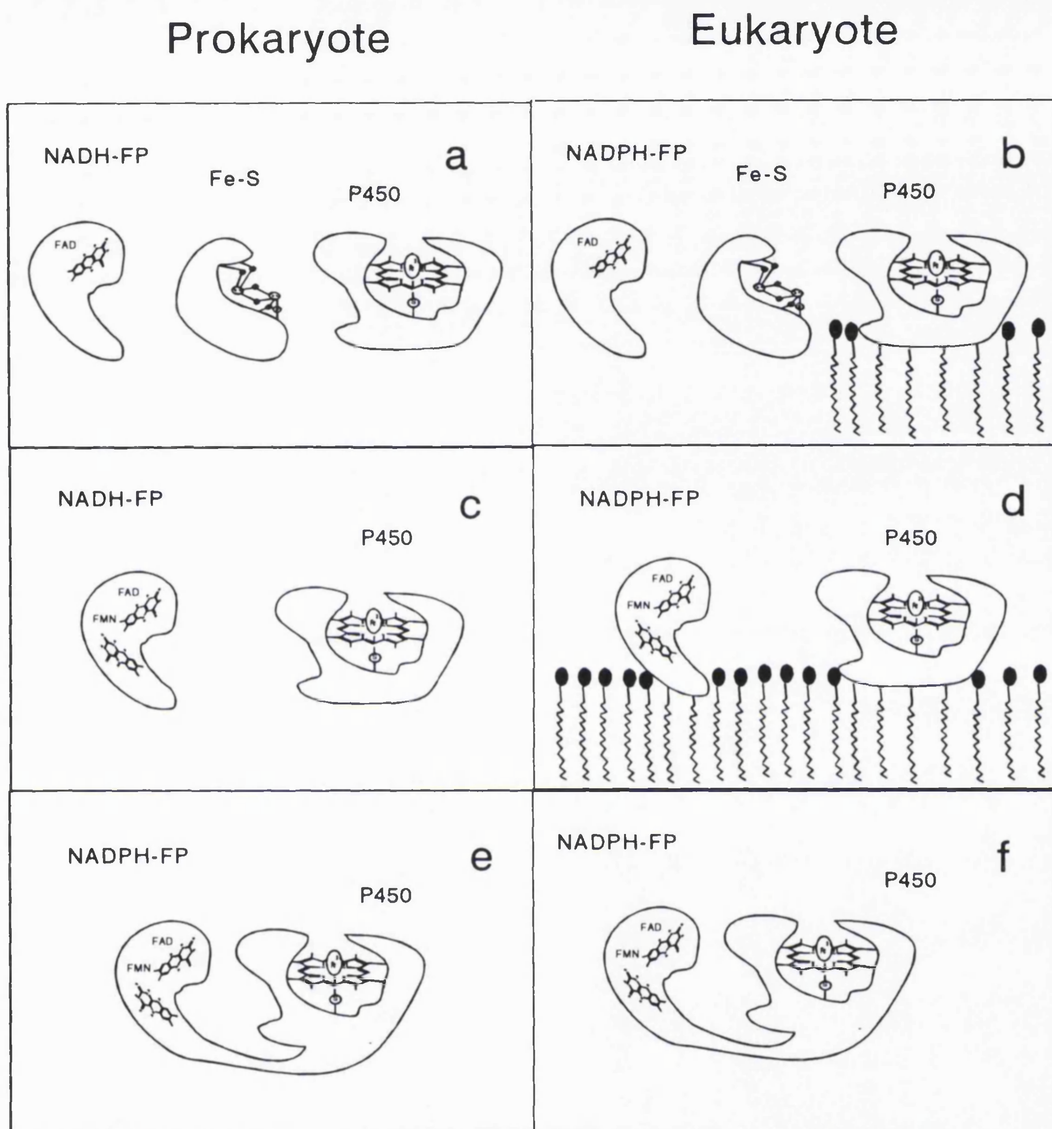
All P450s, however, have a similar mode of action, involving the proton catalyzed splitting of molecular oxygen. The mechanism of the P450 cycle is still under debate, but years of investigations involving ^{18}O-labeling, observed stoichiometry and loss of configuration during oxidation, have lead to a conceptual model of the reactions that occur. The P450 reaction cycle is shown in Figure 1.5b. Initially there was some debate as to whether each electron carrier transferred two electrons (Yasukochi & Masters, 1976; Dignam & Strobel, 1977), but the current opinion is that each carrier transfers a single electron.

1.5.2 Evolution by gene duplication and divergence

What is now called the P450 superfamily encodes numerous enzymes of which over 200 have so far been characterized, from bacteria, yeast, plants and other eukaryotic species. The superfamily comprises 36 gene families, of which 10 include subfamilies and 26 are represented by only one gene (Nelson *et al*, 1993). Some mycoplasma genera do not contain cytochromes at all (Pollack, 1979), yet contain P450s. The P450 enzymes can be considered one of the most ancient respiratory systems. The numerous P450 genes are believed to have arisen from a common sequence by a process of gene duplication and subsequent divergence to generate new enzymes. This process is shown well by the duplication of the human *CYP21* and *CYP11* genes.

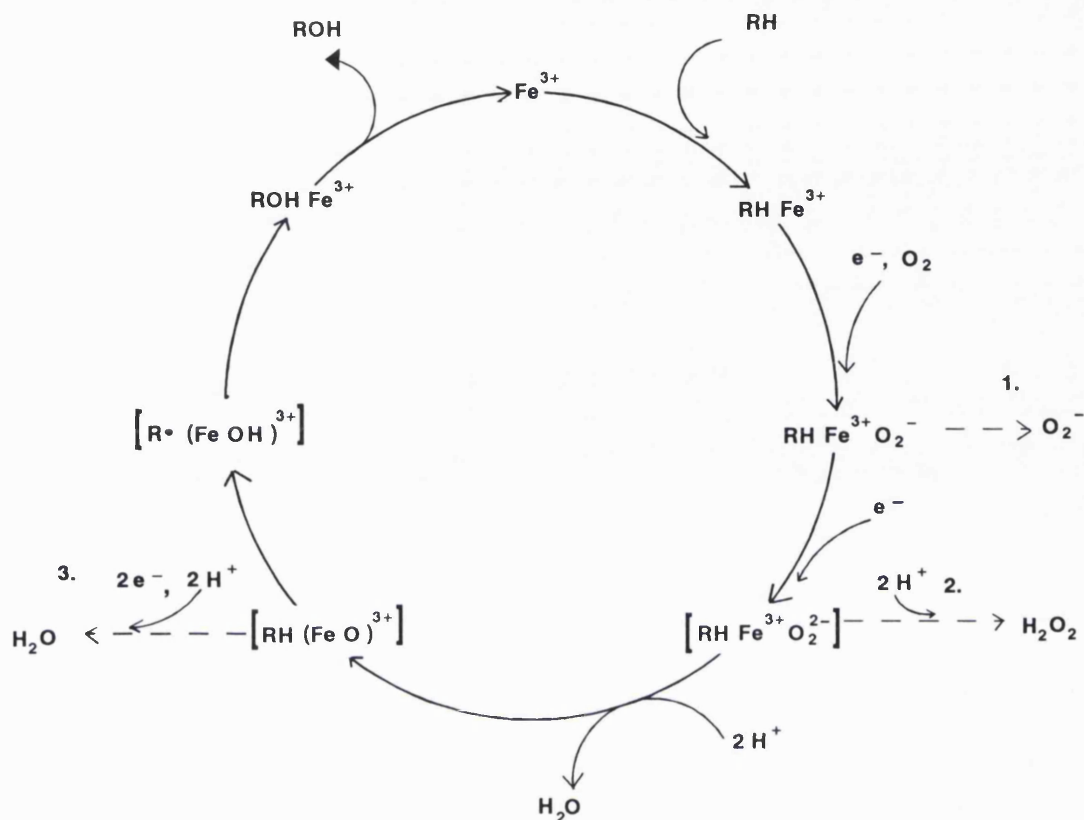
Duplication of genes can occur by unequal crossover during non-homologous recombination or may have occurred by duplication of entire chromosomes in early eukaryotes. Evolutionary studies suggest that the gene duplication that gave rise to the two different classes of P450s may have occurred approximately 1,360 million years ago (Nelson & Strobel, 1987). The *CYP17* gene is believed to be derived from an ancestral gene resembling *CYP21*. The low amino acid homology of these two

Figure 1.5a: P450 enzymes act as terminal oxidases in a number of different electron transfer systems in prokaryotes and eukaryotes.



P450-containing monoxygenase systems fall into six groups: (a) bacterial three component system, e.g. putidaredoxin reductase, putidaredoxin and CYP101 in *Pseudomonas putida*; (b) mitochondrial three-component system, e.g adrenodoxin reductase, adrenodoxin and CYP11B2 in human adrenocortical mitochondria; (c) bacterial two component system, e.g. NADH-dependent P450 reductase and $P450_{sca}$ in *Streptomyces carbophilus* (Serizawa & Matsuoka, 1991); (d) microsomal two component system, e.g. NADPH-dependent P450 reductase and CYP17 in human adrenocortical microsomes; (e) bacterial one component system, e.g. CYP102 in *Bacillus megaterium*; (f) soluble one component P450-like system, e.g. inducible nitric oxide synthase in mouse (White & Marletta, 1992). NADH-FP and NADPH-FP denote NADH- and NADPH-dependent flavoproteins, respectively; Fe-S denotes an iron-sulphur protein.

Figure 1.5b: The proposed P450 reaction cycle (taken from Coon & White, 1980).



The P450 reaction cycle is shown above. RH represents the steroid substrate, ROH the hydroxylated product and Fe^{3+} the iron at the centre of the P450 haem group. Intermediates shown in brackets are hypothetical and have not been directly observed. Reactions 1., 2. and 3., coming off from the cycle as dashed arrows, represent uncoupling. The uncoupling reactions generate superoxide, hydrogen peroxide and excess water. These side reactions compete with the hydroxylation reaction for available electrons.

When RH binds to the P450 the haem can accept electrons. Molecular oxygen binds at the haem after the addition of one electron to the Fe^{3+} iron, to give the complex RH Fe^{3+} $O_2^{\cdot}$. The addition of a second electron produces the complex [RH $Fe^{3+}O_2^{2-}$]. The breakdown of [RH $Fe^{3+}O_2^{2-}$] produces activated oxygen within the [RH $(FeO)^{3+}$] complex, and water. The activated oxygen at the haem abstracts hydrogen from RH, to give [R· $(FeOH)^{3+}$], and the hydroxyl group attacks the carbonyl radical, R·, to give the hydroxylated steroid ROH. The production of activated oxygen is the key point in P450 mediated reactions.

enzymes (28.9%) indicates that the divergence of the two genes is very ancient, and possibly begun around 900 million years ago. A duplication around the same time gave rise to the drug metabolising microsomal P450s, which are related in structure to CYP17 and CYP21 (Nelson & Strobel, 1987). *CYP21* contains 9 small introns, whilst *CYP17* contains only 7 larger introns. This situation has arisen by the loss of intron 1 and 3 from an ancestral form of *CYP17*. The intron/ exon boundaries of *CYP17* and *CYP21* otherwise match and key amino acids are conserved (Picardo-Leonard & Miller, 1987). The selective loss of introns is a characteristic of genetic evolution. It has been suggested that an early ancestral P450 may have had more than 100 mini-exons (Nebert *et al*, 1988). The reason that more ancient genes retain introns may be due to selective pressure to maintain gene structure and may be assisted by transcription coupled repair (Bohr *et al*, 1985; Vos & Waulthier, 1991). This may contrast with newly duplicated genes, which may not initially be transcribed and are certainly non-essential due to the presence of the original gene.

In the case of *CYP17* and *CYP21* duplication has given rise to two genes, under different hormonal control, on separate chromosomes and encoding enzymes with different substrate specificities. A similar process lead to the presence of *CYP11A* on chromosome 15 and an ancestral form of *CYP11B* on chromosome 8.

The high homology of the two *CYP11B* genes suggests that the duplication event that produced them occurred more recently than the one described for the microsomal steroidogenic P450 enzymes. That mice and rats have two different *CYP11B* genes suggests that the *CYP11B* duplication may have occurred before mammalian divergence, around 100 million years ago (Nelson *et al*, 1993). The 98% homology of *CYP21* and *CYP21P* may suggest that this is the most recent duplication event observed.

The differences in sequence between *CYP11B1* and *CYP11B2* are a record of mutations that have occurred over perhaps 100 million years. The fact that the exonic sequences only differ by single base changes suggests that the differences may represent points at which a mutation has occurred in only one of the duplicated genes. This allows analysis of the rate of certain types of mutation within these two sequences to be carried out (see section 6.2.4).

Of greater interest is the identification of amino acids that are vital to structure and function. The presence of so many characterized P450 amino acid sequences allows comparison to detect possible structural motifs and amino acids that may take part in catalysis of the P450 hydroxylation of steroids.

1.5.3 Comparison of P450 enzyme amino acid sequences and structures

Amino acids that are conserved in a large number of P450 enzymes are likely to have specific functions necessary for enzyme activity. Certain amino acid changes, for example lysine to arginine, are relatively conservative in that the two amino acids are very similar in properties, in this case both having long positively charged side chains. Other changes, for example a change of an aspartic acid residue to glycine, can radically alter protein structure and function; aspartic acid can act as a proton donor for catalytic reactions, where glycine cannot, whilst glycine, with only a hydrogen atom as its side chain, can allow rotation of residues linked to it, leading to flexibility of structure and altered conformation. The tertiary structure of an enzyme is the result of interactions between amino acids that can be hundreds of residues apart in the polypeptide chain, yet very close to each other when the protein folds up in three dimensions.

To date only two P450 enzyme structures have been examined by X-ray crystallography: the soluble bacterial P450 enzyme from *Pseudomonas putida*, CYP101 (Poulos *et al*, 1985) and the haem domain of CYP102 from *Bacillus megaterium* (Ravichandran *et al*, 1993). CYP101 is of the same class of P450 as the mitochondrial enzymes, whilst the CYP102 haem domain is of the same class as the microsomal P450s.

The membrane bound mammalian P450 enzymes have proved difficult to crystalize due to their higher degree of hydrophobicity. The amino acid sequence of CYP101 is only 15% homologous to that of mammalian P450s, which makes comparison difficult (Lin *et al*, 1993). Associations with either microsomal or mitochondrial membrane are also likely to alter the folding and positioning of the mammalian P450 enzymes. Nevertheless, common P450 structural motifs appear to be conserved.

In both crystalized enzymes there are 12 main α-helical regions, lettered A-L and at least five ß-pair regions, ß1-ß5 (Poulos *et al*, 1985; Ravichandran *et al*, 1993). CYP101, which is only 414 amino acids in length, is thought to represent the minimal requirements for a P450: hydrophobic substrate binding site, redox partner docking region, haem centre and amino acids arranged for donation of protons to the haem. CYP101 metabolizes camphor, which only contains 11 carbon atoms. The majority of P450s are larger, at around 500 amino acids and have insertions involving larger peripheral loops, extra helical sections and enlarged substrate binding sites (Ravichandran *et al*, 1993) to accomodate larger hydrophobic compounds, such as fatty acids or cholesterol, which has 27 carbon atoms.

As the eukaryotic microsomal P450 enzymes are thought to exist as globular cytosolic proteins, bound to the microsomes only by a hydrophobic anchor and several peripheral hydrophobic contacts, the structural similarity with the soluble bacterial protein CYP102 is presumably very high. Mitochondrial proteins, which are associated with the inner mitochondrial membrane only by peripheral interactions, should also have considerable structural homology with bacterial proteins, particularly in the adrenodoxin and haem binding regions.

Groups working on the structure and function of the P450s have suggested that particular amino acids play key roles. In a review of the evolution of P450 proteins Nelson & Strobel (1987) aligned 34 P450 enzyme amino acid sequences from various mammalian species with the sequence of CYP101. The alignment of CYP101, CYP11A, CYP17 and CYP21 from Nelson & Strobel was used during this study to align CYP11B1 and CYP11B2 amino acid sequences and show which regions are conserved. In this alignment regions of CYP101 and CYP102 that are believed to form secondary α-helices and ß-pairs are highlighted (section 6.4).

1.6 Aims of study

The initial hypothesis of this thesis was that the higher observed mutation rate within the *CYP11B1* gene, relative to that of *CYP17*, was due to non-homologous recombination and gene conversion between the duplicated *CYP11B* genes. In the following study the genetic processes that lead to mutations in the *CYP17*, *CYP11B1* and *CYP11B2* genes were examined.

During the course of the study, mutations were found that would produce enzymes with altered amino acid composition. Determining the effect of these amino acid changes on substrate binding and enzyme action was the second aim of the study. The alignment shown in section 6.4 allowed amino acid changes to be positioned on hypothetical models of the enzymes and provided an insight into the relevance of mutations to enzyme function. By defining which regions or residues of the proteins are essential for their correct function, it is hoped that a clearer idea of their structure and mode of action in producing steroid hormones can be pieced together.

2- Materials and Methods

2.1 Reagents

All reagents used were Analar grade and were obtained from BDH, Lutterworth, Leicestershire, or Sigma Chemical Company Ltd., Poole, Dorset, United Kingdom, unless otherwise stated.

2.1.1 Stock solutions

0.5 M EDTA (pH 8.0): Add 186.1 g disodium ethylenediaminetetra-acetate.2 H_2O to 800 ml of ddH_2O. Stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH (approximately 20 g). Adjust volume to 1 l with ddH_2O, dispense into aliquots and sterilize by autoclaving.

1 M magnesium chloride : Dissolve 203.3 g $MgCl_2.6H_2O$ in 1 l ddH_2O. Autoclave to sterilize.

1 M magnesium sulphate: Dissolve 12 g anhydrous $MgSO_4$ in 100 ml ddH_2O. Autoclave to sterilize.

3 M sodium acetate (pH 5.2): Dissolve 40.8 g sodium acetate.3H_2O in 80 ml ddH_2O. Adjust pH to 5.2 with glacial acetic acid. Adjust volume to 100 ml with H_2O. Dispense into 20 ml aliqouts and sterilize by autoclaving.

5 M sodium chloride: Dissolve 292.2 g NaCl in 800 ml of ddH_2O. Adjust volume to 1 l with ddH_2O. Dispense into aliquots and sterilize by autoclaving.

20xSSC- sodium chloride, sodium citrate buffer: Dissolve 350.6 g NaCl and 176.4 g trisodium citrate in 1500 ml ddH_2O, with stirring on a magnetic stirrer. Adjust pH to 7.0 with a few drops of concentrated citric acid. Adjust volume to 2 l with ddH_2O.

20x SSPE buffer- sodium chloride, sodium dihydrogen phosphate buffer: Dissolve 350.6 g NaCl and 27.6 g $NaH_2PO_4.H_2O$ and 7.4 g EDTA in 800 ml of ddH_2O. Adjust the pH to 7.4 with NaOH (approximately 6.5 ml of a 10 M solution). Adjust the volume to 1 l with ddH_2O. Sterilize by autoclaving.

10% SDS- sodium dodecyl sulphate: Dissolve 292.2 g electrophoresis grade SDS in 900 ml ddH_2O. **Caution:** SDS is irritating to the eyes and respiratory system, when measuring out the powder wear a mask and safety glasses. Heat to 68°C to assist dissolution. Adjust the volume to 1 l with ddH_2O.

10 M sodium hydroxide: In a plastic beaker dissolve 40 g NaOH pellets in 60 ml ddH_2O. Adjust the volume to 100 ml with ddH_2O.

1 M Tris (pH 7.4): Dissolve 13.70 g Tris.HCl and 1.60 g Tris Base in 100 ml ddH_2O. Sterilize by autoclaving.

1 M Tris (pH 7.6): Dissolve 12.70 g Tris.HCl and 2.36 g Tris Base in 100 ml ddH_2O. Sterilize by autoclaving.

1 M Tris (pH 8.0): Dissolve 9.76 g Tris.HCl and 4.60 g Tris Base in 100 ml ddH_2O. Sterilize by autoclaving.

1M Tris (pH 8.3): Dissolve 7.08 g Tris.HCl and 6.68 g Tris Base in 100 ml ddH_2O. Sterilize by autoclaving.

TE buffer- 10mM Tris, 1mM EDTA (pH 8.0): Add 1 ml of 1M Tris (pH 8.0) and 0.2 ml of 0.5 M EDTA (pH 8.0) to 90 ml ddH_2O, mix by inversion then adjust volume to 100 ml. Sterilize by autoclaving.

2.1.2 DNA extraction solutions

Ammonium acetate (7.5 M): Dissolve 28.9 g ammonium acetate in ddH_2O. Adjust volume to 50 ml. Sterilize by filtration through a 0.45 μm filter. Store at room temperature. Reagent has a shelf-life of 6 months.

Guanidine hydrochloride (6 M): Dissolve 57.3 g guanidine hydrochloride in 50 ml ddH_2O, by stirring and heating at 37°C until it dissolves. Adjust volume to 100 ml with ddH_2O. Filter solution through 0.45 μm filter. Store at room temperature. Solution is stable for 6 weeks.

Rapid DNA preparation buffer:

5 ml 1 M KCl	final conc.	50 mM KCl
1 ml 1 M Tris (pH 8.3)		10 mM Tris
0.25 ml 1 M $MgCl_2$		2.5 mM $MgCl_2$
1 ml 1% gelatin		0.1 mg/ ml
0.45 ml NP40		0.45% NP40
0.45 ml Tween 20		0.45% Tween 20
Adjust volume to 100 ml with ddH_2O.		

Autoclave and store frozen in aliquots. Prior to use 3 μl of a 10 mg/ ml proteinase K solution (Promega, Southampton, UK) should be added to each 500 μl aliquot of DNA rapid preparation buffer.

20% sodium sarkosyl: Dissolve 28.91 g in 20 ml ddH_2O, make up to 50 ml with ddH_2O. Filter through a 0.45 μM filter to sterilize.

Sucrose lysis solution: Dissolve 54.7 g sucrose in 400 ml ddH_2O. Add 5 ml Tris.Cl (pH 7.5), 2.5 ml 1 M $MgCl_2$, and 5 ml Triton X-100. Adjust volume to 500 ml with ddH_2O and store at 4°C.

2.1.3 Electrophoresis buffers and associated solutions

Buffer	Working solution	Concentrated Stock (per litre)
Tris-acetate (TAE)	1x: 0.04 M Tris-acetate 0.001 M EDTA	10x: 48.4 g Tris base 11.4 ml glacial acetic acid 20 ml 0.5 M EDTA (pH 8.0)
Tris-borate (TBE)	1x: 0.090 M Tris-borate 0.002 M EDTA	5x: 54 g Tris base 27.5 g boric acid 20 ml 0.5 M EDTA (pH 8.0)
Tris-borate (TBE) for sequencing	1x: 0.100 M Tris-borate 0.001 M EDTA	10x: 121.1 g Tris base 51.4 g boric acid 20 ml 0.5 M EDTA (pH 8.0)

10x Bromophenol blue/ xylene cyanol loading buffer: Dissolve 0.125 g bromophenol blue and 0.125 g xylene cyanol in 30 ml ddH_2O. Add 12.5 g Ficoll (Type 400; Pharmacia, Uppsala, Sweden), sprinkling small amounts on top of the solution and stirring on a magnetic stirrer until all the Ficoll has dissolved. Adjust volume to 50 ml with ddH_2O and aliquot into 1 ml amounts. Store at room temperature, or freeze at -20°C for later use.

Formamide/ bromophenol blue loading buffer:

	final concentration:
19 ml formamide	95% formamide
0.8 ml 0.5 M EDTA (pH 8.0)	20 mM EDTA
0.2 ml 5% bromophenol blue solution	0.05 % bromophenol blue

Ethidium bromide: Add 0.25 g of ethidium bromide to 25 ml ddH_2O. Stir on a magnetic stirrer for several hours to ensure that the dye has dissolved. Wrap the container in aluminium foil and store at room temperature. The ethidium bromide solution is used in agarose gels at a concentration of 0.5 μg/ ml, i.e 5 μl stock solution in 100 ml agarose gel.

1% nitric acid (used in silver staining DNA on polyacrylamide gels): Add 28.5 ml concentrated nitric acid to 1900 ml of ddH_2O in a fume hood. Adjust volume to 2 l with ddH_2O. **Caution:** Concentrated nitric acid is highly corrosive. Wear gloves, protective clothing and safety glasses when handling.

Silver nitrate solution (0.012 M $AgNO_3$) used in silver staining: Dissolve 4.1 g $AgNO_3$ in 2 l of ddH_2O.

Sodium carbonate/ formaldehyde silver stain developing solution: Dissolve 178.2 g Na_2CO_3 in 2 l ddH_2O. In a fume hood add 3 ml of formalin (purchased as a 40% formaldehyde solution in ddH_2O, from Sigma) and stir on a magnetic stirrer. Adjust volume to 6 l with ddH_2O, to give final concentration of 0.28 M Na_2CO_3, 0.02% formaldehyde.

Silanising solution (5% dimethyldichlorosilane in chloroform): In a fume hood measure out 95 ml of chloroform into a glass bottle that can be well sealed. Using a glass pipette and bulb add 5 ml of dimethyldichlorosilane to the chloroform and gently rinse the pipette out several times with the chloroform. Cap the bottle securely and carefully invert to mix the solution. Store at 4°C.

Equilibrated phenol: Phenol is used in extractions to remove proteins and agarose polymers from DNA bands cut out from agarose gels. Before use it must be equilibrated to a pH > 7.8, as DNA will partition into phenol at acid pH. This is carried out by equilibration with 1.0 M and 0.1 M Tris.Cl (pH 8.0) according to the method of Sambrook *et al* (1989a).

Phenol: chloroform : isoamyl alcohol (25:24:1): A mixture consisting of equal parts of equilibrated phenol and chloroform: isoamyl alcohol (24:1) is frequently used to remove proteins from preparations of nucleic acids. The chloroform denatures proteins and facilitates the separation of aqueous and organic phases, and the isoamyl alcohol reduces foaming during extraction. Neither chloroform nor isoamyl alcohol requires treatment before use.

2.1.4 Southern blotting solutions

Alkali blotting denaturing and transfer solution (0.4 M NaOH): Dissolve 32 g sodium hydroxide pellets in 1500 ml ddH_2O, stirring on a magnetic stirrer, then adjust volume to 2 l to give a final concentration of 0.4 M NaOH.

100x Denhardt's solution: 10 g Ficoll (Type 400, Pharmacia, Uppsala, Sweden)
10g polyvinylpyrrolidone
10g BSA (Fraction V: Sigma)
Make up to 500 ml with ddH_2O.

DNA hybridisation buffer:

25 ml 20x SSPE	final concentration	5x SSPE
5 ml 100x Denhardt's		5x Denhardt's
5 ml 10% SDS		0.5% SDS
Make up to 100 ml ddH_2O.		

0.25 M HCl: Wearing safety glasses and gloves add 43 ml concentrated hydrochloric acid to 1 l ddH_2O in fume cupboard, while stirring gently on a magnetic stirrer. Adjust volume to 2 l.

Solution O: 1.25 M Tris (pH 8.0), 0.125 M $MgCl_2$.

Solution A: 1 ml solution O
18 μl 2ß-mercaptoethanol
5 μl 0.1 M dATP
5 μl 0.1 M dGTP
5 μl 0.1 M dTTP

Solution B: 2 M Hepes NaOH (pH 6.6)

Solution C: Hexadeoxyribonucleotides 4.22 mg/ ml (Pharmacia, Uppsala, Sweden)

OLB buffer: Made from solution A, solution B and solution C in the ratio of 100: 250:150. Store at -20°C.

DNA denaturing solution (S1): Dissolve 175.3 g NaCl in 1500 ml ddH_2O by stirring on a magnetic stirrer, add 40 g NaOH pellets and stir until dissolved. Adjust volume to 2 l, to give final solution 1.5 M NaCl, 0.5 M NaOH.

Neutralizing solution (S2): Dissolve 175.3 g NaCl in 1500 ml ddH_2O by stirring on a magnetic stirrer, add 121 g Tris base, dissolve by stirring, then add 4 ml 0.5 M EDTA (pH 8.0) solution. Adjust pH to 7.2 with approximately 60 ml concentrated hydrochloric acid. Adjust volume to 2 l, to give final solution 1.5 M NaCl, 0.5 M Tris, 1 mM EDTA (pH 7.2).

Salmon sperm DNA(10 mg/ ml): Add 1 g of salmon sperm DNA to 100 ml ddH_2O. Stir on a magnetic stirrer for 2-4 hours at room temperature to help DNA dissolve. The DNA is sheared by passing it 12 times rapidly through a 17-gauge hypodermic needle, sonicated and stored at -20°C. Just before use the solution is bolied for 5 minutes and chilled quickly on ice. Denatured, fragmented salmon sperm DNA should be used at a concentration of 100 μg/ ml in prehybridisation solutions.

2xSSC: Add 100 ml 20x SSC solution to 800 ml ddH_2O, then adjust final volume to 1 l.

2.1.5 Bacterial growth media and solutions for gene library screening

Ampicillin 100 mg/ ml: Dissolve 500 mg of ampicillin (Sigma) in 5 ml sterile ddH_2O by inversion within a sterile 15 ml tube. Dispense into 500 μl aliquots and store at -20°C. Dilute 1 in 1000 when selecting ampicillin resistant bacteria in LB-ampicillin broth or on LB-ampicillin agar plates.

2% 5-Bromo-4-chloro-3-indolyl-ß-D galactoside (X-gal) in dimethylformamide: Weigh out 20 mg of X-gal (Promega, Southampton, UK) into a 1.5 ml eppendorf. In a fume hood add 1 ml dimethylformamide and invert to dissolve X-gal. X-gal powder and this 2% stock solution should be stored at -20°C in darkness. X-gal is used at a final concentration of 0.01% in LB-Xgal agar plates.

50mM Calcium chloride: Dissolve 6.75 g $CaCl_2.6H_2O$ in 400 ml of ddH_2O then make up to 500 ml. Sterilize by autoclaving.

100 mM Isopropylthio-ß-D-galactoside (IPTG): Weigh out 23 mg of IPTG (Promega, Southampton, UK) into a 1.5 ml eppendorf. Add 1 ml of ddH_2O and dissolve by inversion. This 100 mM stock solution is diluted 1 in 500 when making LB-ampicillin, IPTG, X-gal agar plates. IPTG powder and solution should be stored at -20°C.

Luria-Bertani (LB) broth: Dissolve 10 g bactotryptone, 5 g yeast extract and 10 g NaCl in 800 ml ddH_2O with stirring on a magnetic stirrer. Adjust pH to 7.5 with approximatley 10 drops of 10 M NaOH solution. Adjust volume to 1 l and autoclave to sterilize.

LB-agar for bacterial plates: To 500 ml of LB broth add 7.5 g bactoagar (Difco, Detroit, USA). Autoclave to sterilize, swirl to mix following autoclaving, allow to cool to below 50°C and pour 25 ml per 90 mm diameter plate. LB-agar ampicillin plates are made as above except that ampicillin at a final concentration of 100 μg/ ml is added to the autoclaved LB-agar after it has cooled to below 50°C. LB-ampicillin, IPTG, X-gal agar plates, for ß-galactosidase blue-white selection, contain 100 μg/ ml

ampicillin, 0.2 mM IPTG, 0.01% X-gal. LB-agar $MgSO_4$ plates are as above but contain 10 mM $MgSO_4$.

LB-top agar: As for LB-agar, but using 0.7 g bactoagar per 100 ml LB-broth.

NZY medium: Dissolve 10g NZ amine (casein hydrolysate), 5g NaCl, 5g yeast extract and 2g $MgSO_4.7H_2O$ in 900 ml ddH_2O. Adjust pH to 7.0 with 10 M NaOH, make up to 1 l with ddH_2O, then sterilize by autoclaving.

NZY-agar: Prepare NZY medium as above. To 500 ml of medium add 7.5 g of bactoagar. Sterilize by autoclaving, allow to cool below 50°C and pour into plates.

3 M potassium acetate (pH 5.5): Dissolve 294.45 g potassium acetate in 500 ml ddH_2O. Adjust pH to 5.5 with glacial acetic acid (approximately 110 ml). Adjust volume to 1 l with ddH_2O. Sterilize by autoclaving.

Proteinase K buffer: Make 3.2 ml of 5 M NaCl and 1.8 ml of 0.5 M EDTA (pH 8.0) up to 100 ml with sterile ddH_2O invert to mix and use immediately.

SDS/ NaOH lysis solution:

2 ml 10 M NaOH	final concentration	0.2 M NaOH
10 ml 10% SDS		1% SDS
Make up to 100 ml.		

Phage dilution (SM) buffer: Dissolve 5.8 g NaCl and 2.0 g $MgSO_4$ in 800 ml ddH_2O. Add 50 ml 1 M Tris, (pH 7.5) and 5 ml of 2% gelatin. Adjust volume to 1 l, mix well by inversion and autoclave to sterilize.

STE buffer:

20 ml 5 M NaCl	final concentration	0.1 M NaCl
5 ml 1 M Tris (pH 8.0)		10 mM Tris (pH 8.0)
1 ml 0.5 M EDTA (pH 8.0)		1 mM EDTA (pH 8.0)

Make up to 500 ml with ddH_2O.

Terrific broth (TB): Dissolve 5 g NaCl and 10 g bactotryptone in 800 ml ddH_2O. Adjust pH to 7.4 with 10 M NaOH solution. Adjust volume to 1 l with ddH_2O and sterilize by autoclaving. (For bacteriophage transfection work add sterile $MgSO_4$ to a final concentration of 10 mM and sterile maltose to a final concentration of 0.2%).

50 mM Tris (pH 8.0), 10 mM EDTA containing 100 μg/ ml RNase A: Dissolve 6.055 g Tris base, 3.722 g EDTA.2H_2O in 800 ml ddH_2O. Adjust pH to 8.0 with hydrochloric acid. Allow the solution to cool to room temperature before making the final adjustments to the pH. Adjust volume to 1 l with ddH_2O and sterilize by autoclaving. When solution is cooled to room temperature, add RNase A (Promega, Southampton, UK) to a concentration of 100 μg/ ml to a single 100 ml aliquot and store at 4°C.

2.1.6 Sequencing solutions

The following solutions are provided in the Sequenase version 2.0 kit™ (United States Biochemicals, Cambridge Bioscience, Cambridge, UK).

Manganese buffer: (5x concentrate)

0.15 M sodium isocitrate
0.1 M $MnCl_2$

Sequenase buffer: (5x concentrate)

200 mM Tris.Cl
100 mM $MgCl_2$
250 mM NaCl

Primer (-40 Universal primer): 5'-GTTTTCCCAGTCACGAC-3' 0.5 pmol/ μl

Dithiothreitol (0.1 M)

Labelling mix (dGTP): (5x concentrate)

7.5 μM dGTP
7.5 μM dCTP
7.5 μM dTTP

ddG Termination mix: 80 μM : dGTP, dATP, dTTP, dCTP
8 μM : ddGTP
50 mM : NaCl

ddA Termination mix : As for ddG mix except 8 μM ddGTP is replaced by 8 μM ddATP.

ddT Termination mix : As for ddG mix except 8 μM ddGTP is replaced by 8 μM ddTTP.

ddC Termination mix : As for ddG mix except 8 μM ddGTP is replaced by 8 μM ddCTP.

Stop solution: 95% formamide
20 mM EDTA
0.05% bromophenol blue
0.05% xylene cyanol

Enzyme dilution buffer: 10 mM Tris (pH 7.5)
5 mM dithiothreitol
0.5 mg/ ml BSA

Sequenase version 2.0 T7 DNA polymerase: 13 units/ μl
20 mM KPO_4 (pH 7.4)
1 mM DTT
0.1 mM EDTA
50% glycerol

Sequencing gel (6% polyacrylamide 19:1 acrylamide: bisacrylamide): To 20 g of sequencing grade urea, in a glass beaker, add 18 ml of ddH_2O, 4 ml of 10xTBE for sequencing buffer and 6 ml Accugel (National Diagnostics, Aylesbury, UK). Swirl to mix then incubate at 37°C for 10 minutes to assist dissolution of the urea. Swirl gently until all urea has dissolved. Transfer to a 50 ml falcon tube and add 400 μl of a 10% ammonium persulphate solution. Invert gently to mix, then add 20 μl N,N,N',N'-tetramethylethylenediamine (TEMED). Mix, then pour between taped 20x30 cm glass plates, containing 0.4 mm spacers. Allow to set for at least 1 hour.

2.2 Methods

2.2.1 DNA extraction from blood leucocytes: (Jeanpierre, 1989)

The following protocol for DNA extraction from whole blood was the standard method employed. Yields of DNA in the range of 0.1-1.0 mg/ ml (50-500 μg total DNA) were routinely obtained from 5 or 10 ml of whole blood. This DNA was used for Southern blotting and as template for PCR reactions.

10 ml whole blood was collected into potassium EDTA Vacutainer tubes and stored at -20°C. Details of each patient were noted, and a number assigned to each sample. Samples were thawed and blood transferred to 50 ml falcon tubes. Tubes were filled to 40 ml mark with ice-cold water, inverted to mix, then spun at 3000 rpm for 20 minutes at 4°C. The supernatant was poured off into 1% sodium hypochlorite solution.

The pelleted cells were resuspended in 25 ml of 0.1% NP40 solution using a vortex mixer to break up the pellets. Nuclei were pelleted by spinning at 3000 rpm for 20 minutes at 4°C. The supernatant of cellular debris was discarded into sodium hypochlorite.

The nuclei were vortexed hard in a solution comprised of 3.5 ml 6 M guanidine hydrochloride and 0.25 ml 7.5 M ammonium acetate, to solubilize the pellet. Digestion of protein in the nuclei was carried out by adding 0.25% sodium sarkosyl, 75 μl of a 10 mg/ ml proteinase K solution and incubating at 60°C for 1.5 hours.

The samples were then cooled to room temperature and 8.5 ml absolute ethanol added to precipitate the extracted DNA, which appears as a small white mass of threads. The DNA was spooled out on a glass pipette, allowed to dry briefly and added to 1 ml TE buffer in a bijou tube. The DNA was dissolved overnight on a rotating wheel. The samples of DNA were reprecipitated by adding one tenth volume of 3 M sodium acetate (pH 5.2) and 2 volumes of ice cold absolute ethanol.

The DNA was spooled out once more, allowed to dry and dissolved overnight in 0.5 ml TE buffer on a rotary wheel. Samples were stored at -20°C.

The concentration of DNA was measured by adding 10 μl of DNA solution to 990 μl of ddH_2O and then recording the absorbance at 260 nm on a spectrophotometer. DNA concentration was deduced by the following formula:

Concentration of DNA in μg/ ml = Absorbance at 260nm x dilution x 50 *

* Derived from 50 μg/ml DNA solution giving an absorbance reading of 1 at 260 nm (Sambrook *et al,* 1989b).

2.2.2 Rapid DNA preparation from 0.5 ml of whole blood (Ehrlich, 1989)

A rapid procedure was used to prepare DNA for PCR. This method required only 0.5 ml of blood and was therefore a useful source of genomic DNA when there was little blood available from a patient, for instance with very young children. The DNA prepared by this method was not suitable for restriction endonuclease digestion.

0.5 ml blood was mixed with 0.5 ml sucrose lysis buffer in a 1.5 ml eppendorf tube, which was then spun at 13,000 rpm for 20 seconds at room temperature. The supernatant was decanted into 1% sodium hypochlorite solution and the pellet resuspended, using a vortex mixer, in 1 ml sucrose lysis solution. The centrifugation and resuspension in sucrose lysis solution were repeated twice more or until the pellet was free of haemoglobin.

The sample was spun at 13,000 rpm for 20 seconds, the supernatant removed and the pellet resuspended in 0.5 ml DNA rapid preparation buffer and 3μl of 10 mg/ ml proteinase K (Promega, Southampton, UK). The resuspended pellets were digested at 55°C for 1 hour. If the pellet had not digested a further 3 μl of 10 mg/ ml proteinase K was added and the samples reincubated for another 60 minutes at 55°C. The samples were incubated at 90°C for 10 minutes to inactivate the proteinase K enzyme and stored at -20°C.

2.2.3 DNA preparation from buccal cells (Lench *et al*, 1988)

Obtaining DNA from mouthwashes proved useful in that samples could be taken from whole families and young children without the need for an invasive procedure or staff trained and licensed to take blood. The samples proved adequate for PCR, although only small quantities of DNA were obtained.

Subjects were asked to rinse the mouth vigorously with approximately 15 ml of ordinary drinking water and spit the solution into a sterile 20 ml tube. Children were encouraged to use straws and suck the water in and out of the collection tube. The mouthwashes were spun at 3000 rpm for 10 minutes at room temperature to pellet the buccal cells and the supernatant discarded into 1% sodium hypochlorite solution for disinfection. The pellets were resuspended in 250 μl of 50 mM NaOH and transferred to 1.5 ml eppendorf tubes. The samples were heated at 100°C for 20 minutes in a boiling water bath, then allowed to cool. The samples were adjusted to pH 7.0 with 50 mM hydrochloric acid (approximately 200 μl). The pH was tested by removing 20 μl aliquots and spotting onto pH paper. Once neutralized the DNA samples were diluted 1 in 2 with sterile ddH_2O and 10 μl of this diluted solution used in a 25 μl PCR reaction.

2.2.4 Polymerase chain reaction amplification

PCR reactions were carried out in 25 μl volumes. Primer pair master mixes were prepared with the following standard composition:

		Mastermix	Final concentration in 25 μl PCR reaction
10x Taq Polymerase Buffer	:	2.5x	1x
2.5 mM dNTPs (Promega)	:	500 μM	200 μM
25 mM $MgCl_2$	:	3.75 μM	1.5 μM
5' primer	:	2.50 μM	1.0 μM
3' primer	:	2.50 μM	1.0 μM

The Promega 10x *Taq* polymerase buffer consists of 500 mM KCl, 100 mM Tris (pH 9.0), 1% Triton X100 and Promega *Taq* DNA polymerase is supplied at a concentration of 5 U/ μl in storage buffer (20 mM Tris (pH 8.0), 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol, 0.5% NP40, 0.5% Tween 20).

10 μl of the above master mix were added to sterile ddH_2O. The sample was overlayed with 2 drops sterile mineral oil and 250 ng of genomic leucocyte DNA added under the oil to give a final volume of 20 μl. Samples were placed on a thermal cycler (Hybaid Ltd., Teddington, UK), and heated at 98°C for 5 minutes to denature DNA. As the thermal cycler cooled below 90°C 0.5 U of *Taq* polymerase enzyme (Promega, Southampton, UK) in 5 μl of 1x*Taq* polymerase buffer was added to each sample and the thermal cycler cooled to 58°C. The samples then underwent 35 cycles of 94°C for 30s, 58°C for 30s, and 72°C for 1 minute. A final step of 72°C for 5 minutes was incorporated into the thermal cycler programme to complete extension of products.

Variation of standard PCR conditions

Titration of magnesium ion concentration from 0-2.5 mM was carried out with primers that gave multiple products or poor yields. The annealing temperature was reduced or increased depending on the Tm value of primers: the temperature in °C at which 50% of the primer molecules are no longer bound to complementary DNA. Tm values were calculated approximately by the Wallace formula:

$$Tm = 4x(\text{no. of G and C bases in primer}) + 2x(\text{no. of A and T bases in primer})$$

A table of primers and Tm values is given in Appendix 1. PCR reaction annealing temperatures were generally carried out around 5-10°C below the average Tm value for the primer pair.

For larger PCR products, up to 2.0 or 3.0 kb in length, the extension time at 72°C was increased from 1 minute to 2 minutes.

Quantification of PCR DNA product

The concentration of purified PCR product was deduced by electrophoresis of a 1 μl aliquot alongside 150 ng of ϕX174 *Hae*III DNA markers (Promega, Southampton, UK) on a 0.8% agarose minigel containing ethidium bromide. The fluorescent intensity of the PCR product band was compared to a band of similar size in the marker ladder and the approximate concentration of PCR product calculated by the formula below:

$$\text{PCR product conc. (ng/}\mu\text{l)} = X \times \frac{\text{Size of comparative marker band (bp)}}{\text{Total size of } \phi \text{ X174 (bp)}} \times 150 \text{ ng}$$

Where X= comparative fluorescent intensity of PCR product to marker band of similar size.

As the largest band of a ϕX174 *Hae*III digest is 1,353 bp, PCR products larger than 2 kb were quantified in the same way, but using lambda *Hind*III markers (Promega,

Southampton, UK) and substituting the values for lambda into the equation above. Table 2.2.4 gives the comparative sizes of DNA marker bands for ϕX174 *Hae*III and lambda *Hind*III and the corresponding mass of each fragment (in ng) in 150 ng total marker DNA.

Table 2.2.4: DNA size markers and relative amounts of DNA within each marker band.

Size of Fragment (bp)	Fraction of total	Mass of fragment in 150 ng total marker DNA (ng)
lambda *Hind*III		
23,130	0.476	71.4
9,416	0.194	29.1
6,557	0.135	20.3
4,361	0.090	13.5
2,322	0.048	7.2
2,027	0.041	6.2
564	0.012	1.8
125	0.002	0.3
Total: 48,502		
ϕX174 *Hae*III		
1,353	0.251	37.7
1,078	0.200	30.0
872	0.161	24.2
603	0.112	16.8
310	0.057	8.6
281	0.052	7.8
271	0.050	7.5
234	0.043	6.5
194	0.036	5.4
118	0.022	3.3
72	0.013	2.0
Total: 5,386		

2.2.5 Restriction endonuclease digestion

The ability of restriction endonuclease enzymes to cut double stranded DNA at specific sequences was used in this project to look for large insertions or deletions in the *CYP17* or *CYP11B* genes, which might lead to a change in the size of expected restriction fragments. These enzymes were also used for the preparation of DNA for cloning and confirmation of some mutations which affected restriction endonuclease recognition sites.

Procedure for restriction endonuclease digestion:

For Southern blot analysis, 10 μg of genomic DNA in a volume of 26 μl sterile ddH_2O was prepared. For PCR product digestion 250 ng in 8.5 μl sterile ddH_2O was used. If necessary DNA solutions were concentrated by means of precipitation with 3 M sodium acetate and ethanol, as detailed in section 2.2.1.

Digestion mix:	Genomic DNA	PCR products
DNA	26 μl	8.5 μl
10x buffer	3 μl	1.0 μl
Endonculease	1 μl (10 U)	0.5 μl (5 U)
	30 μl total volume	10 μl total volume

Restriction endonucleases used in this study were obtained from Promega, Southampton, UK. Enzymes are supplied with a 10x restriction enzyme digestion buffer by most manufacturers. If multiple samples were digested a master mix of enzyme and 10x buffer was prepared just before digestion.

Digests were incubated overnight (16 hours) at optimum temperature for enzyme digestion. Conditions for incubation were obtained from the manufacturers specifications, but generally incubations were at 37°C.

2.2.6 Silver staining of DNA in polyacrylamide gels (Merrill *et al*, 1981).

When a mutation was located which altered a restriction endonculease site, the region around the mutation was amplified by PCR and the DNA product digested with the appropriate enzyme in a total volume of 10 μl, as detailed in section 2.2.5. Normal control PCR reactions and digests were prepared at the same time.

A non-denaturing 10% polyacrylamide gel was then prepared. 10 ml of Accugel (National Diagnostics, Aylesbury, UK), 4 ml of 1x TBE for sequencing and 26 ml ddH_2O were combined in a 50 ml falcon tube and mixed by inversion. 400 μl of 10% ammonium persulphate were added and the solution mixed by inversion once more. 20 μl of TEMED were then added, the solution mixed immediately by inversion and the gel poured between 20x20 cm taped glass plates separated by 0.4 mm spacers. A crenellated comb was pushed into the gel before it had set to produce wells. Accugel contains a 19:1 mixture of acrylamide:bisacrylamide at a concentration of 40% total acrylamide/ bisacrylamide.

Each 10 μl digestion was mixed with 4 μl of 95% formamide/ bromophenol blue loading buffer and 2 μl of 50% glycerol solution. The samples were electrophoresed on the 10% polyacrylamide gel for 2-4 hours at 250 V, 8 mA, alongside ϕX174 *Hae*III DNA size markers, which had been prepared with the same formamide/ bromophenol blue loading buffer and glycerol solution. After electrophoresis the glass plates were prised apart, the gel left attached to one plate and subjected to the following staining procedure:

The gel was immersed, still on the glass plate, in 10% ethanol for 5 minutes to immobilize the DNA in the gel.

The 10% ethanol was drained off the gel and the gel immersed in 1% nitric acid for 5 minutes to oxidize the DNA.

The gel was removed from nitric acid and immersed in ddH_2O for about 20 seconds, to remove excess nitric acid.

The ddH_2O was drained off the gel and the gel immersed in 0.012 M $AgNO_3$ solution for 20 minutes. Silver ions associated with the oxidized DNA.

The $AgNO_3$ solution was drained off and the gel immersed for 20 seconds in ddH_2O to remove excess $AgNO_3$.

The silver staining was developed by immersing gel in sodium carbonate/ formaldehyde developing solution. Silver ions were reduced to silver atoms, which gave staining of the DNA. The developing solution was changed when it darkened and developing continued until DNA bands appeared on the gel.

The silver staining procedure given was able to visualize as little as 1 ng of DNA and the above 10% polyacrylamide gel was capable of resolving differences in length of double stranded DNA down to 5 bp in an 80 bp fragment. Typical sizes of restriction digestion fragments analyzed were from 60 to 300 bp.

2.2.7 Purification of DNA from agarose

DNA produced by PCR was used for cloning and direct sequencing. To separate the desired PCR product from excess primers and non-specific PCR products, the DNA was electrophoresed on 0.8% agarose gels, containing ethidium bromide at a concentration of 0.5 μg/ml, in 1x TBE or 1x TAE electrophoresis buffer. The required PCR band was excised from the gel and purified away from the agarose, by the following methods.

Method 1: Electroelution of DNA from agarose gel

Electroelution was used to purify large fragments of DNA, in the order of 2.0-7.0 kb in length, as the method is relatively gentle and less shearing of DNA occurs compared to the other two methods described.

A plasticine bridge was set up across an electrophoresis tank. Each side was filled with 0.5x TBE buffer. The plasticine was just high enough to physically separate the buffer on the two sides. The electroelution apparatus was then prepared as follows. Lengths of dialysis tubing were cut and slit up the side, then boiled in a 2% sodium bicarbonate/ 1 mM EDTA (pH 8.0) solution for 10 minutes. The tubing was rinsed thoroughly in distilled water and boiled for 10 minutes in 1 mM EDTA. The tubing was allowed to cool then stored in a 50% ethanol solution at 4°C.

Small pieces of treated dialysis tubing were cut and clipped over either end of two chambers of the electroelution apparatus, with the plastic rings provided. The apparatus was then placed over the plasticine bridge, so that the large chamber was on the anode side and smaller chamber on the cathode side. The chambers of the apparatus were filled with 0.5x TBE. Slices of agarose gel containing DNA were placed into the large chamber, then a 25 V potential applied across the tank for 2 hours. The DNA migrated out of the gel towards the cathode and collected on the dialysis membrane of the smaller chamber. After 2 hours the current was reversed for 2 seconds to allow the DNA to move off the dialysis membrane. The buffer in the electroelution apparatus was then carefully removed with a pasteur pipette, leaving

around 300 μl in the small chamber. This buffer, containing the eluted DNA, was collected into a 1.5 ml eppendorf. The DNA was concentrated by precipitation.

Method 2: Spinning DNA in agarose through silanised glass wool

To purify small pieces of DNA, less than 1 kb, from agarose the following method of spinning through glass wool was used. Larger fragments of DNA may undergo shearing during this process and may stick to the silanised glass wool, reducing the yield obtained.

Agarose slices containing DNA were cut into 2 mm square pieces with a sterile scalpel blade. The bottom of a 0.5 ml eppendorf was pierced with a hypodermic needle to give a small hole just less than 1 mm diameter. A small amount of silanised glass wool was placed in the bottom of the eppendorf, to cover the hole. The agarose gel pieces were placed into the eppendorf. The cap was cut off of a 1.5 ml eppendorf and the 0.5 ml eppendorf placed inside. The 0.5 ml tube was spun within the 1.5 ml tube at 13,000 rpm for 2 minutes at room temperature. The centrifugation compressed the gel, which liquified and passed through the glass wool into the bottom tube. The liquified gel was transferred to a 1.5 ml eppendorf. Centrifugation was repeated until all agarose had been liquified. Some agarose remained as a compressed mass in the glass wool and was discarded. The liquified agarose was extracted once with an equal volume of equilibrated phenol and once with an equal volume of chloroform. The aqueous layer was removed and the DNA precipitated with a tenth volume of 3 M sodium acetate (pH 5.2) and 2 volumes absolute ethanol.

Method 3: Purifying DNA from agarose with glass milk (Geneclean™)

Glass milk (Geneclean) was used to purify DNA from agarose, when fragments were larger than 0.5 kb and clean DNA was required for cloning or direct sequencing. Fragments smaller than 0.5 kb remain attached to fine glass particles and are lost during the centrifugation steps of the purification process. Large DNA fragments, above 5 kb, should not be purified by this method as the process can lead to shearing. Purification was carried out as described in the Geneclean protocol using the Geneclean kit (Stratech, Luton, UK). The best results were obtained using 2% NuSieve low melting point agarose (Flowgen, Sittingbourne, UK) made up in 1x TAE. TBE buffer can reduce the binding of DNA to glass milk.

Around 1.0 cm^3 agarose gel was dissolved in 3 ml 3 M sodium iodide solution, with heating at 50°C for 5 minutes. This solution was then mixed well with 5 μl of glass milk slurry, first by inversion then by rotating for 1 hour on a rotary wheel at 4°C, during which time the DNA adheres to the glass particles. Centrifugation at 13,000 rpm for 10 seconds pelleted the glass particles. The pellet was resuspended in 0.5 ml New Wash (a solution supplied in the Geneclean kit), centrifuged again at 13,000 rpm to form a pellet and the supernatant removed. Two more additional wash steps were carried out to clean away agarose from the glass particles and DNA. The final New Wash was removed after centrifugation and the pellet resuspended in 10 μl of ddH_2O. DNA was removed from the glass particles by heating at 55°C for 5 minutes. The glass particles were pelleted by centrifugation at 13,000 rpm for 30 seconds and the supernatant containing DNA transferred to a 1.5 ml eppendorf.

2.2.8 Southern blot analysis

Radioactive labelling of DNA fragments

DNA fragments were purified from agarose gel by spinning through silanised glass wool, as described in section 2.2.7. In a labelled, 1.5 ml eppendorf, 100 ng of DNA was made up to 35 μl with ddH_2O. The DNA solution was boiled for 5 minutes to give single stranded DNA. The sample was spun briefly and placed on ice. When cool the following reagents were added:

BSA (10 mg/ ml)	2 μl
OLB buffer	10 μl
6 U Klenow polymerase (Gibco, BRL)	1 μl
^{32}P-dCTP (10 μCi/ μl, 3000 Ci/ mmol)	3 μl

The reaction mixture was incubated at 37°C for 1 hour or overnight at room temperature. The reaction was stopped by adding 60 μl 5 mM EDTA. The probe was purified away from unincorporated radioactive nucleotides on a column of Sephadex G50 (Pharmacia, Uppsala, Sweden). The probe was stored in a 1.5 ml eppendorf at -20 °C. Before use the probe was thawed, and the tube containing it placed in a boiling water bath for 5 minutes to make single stranded DNA for hybridization.

Alkali Southern blotting using Hybond N^+

Approximately 10 μg of genomic DNA for each subject was digested overnight in a total volume of 30 μl, with an appropriate restriction endonuclease enzyme that would cut at the locus of the gene of interest (for digest conditions see section 2.2.5). After incubation 3 μl of 10x bromophenol blue/ xylene cyanol loading buffer was added to each sample. Electrophoresis was carried out on a 20x20 cm, 0.8% agarose gel in 1xTAE and containing ethidium bromide at a concentration of 0.5 μg/ ml, which was run overnight at 46 V and 40 mA in 1xTAE buffer. Bacteriophage lamda *Hind* III DNA size markers (0.5-23 kb) were loaded on the same gel to give an indication of size. The gel was examined on a UV transilluminator and the distance

migrated by the lambda standards measured. The gel above the wells was cut away with a scalpel, so that the top edge was level with the bottom of the wells and the length and width of gel to be blotted measured. Exposure of the DNA to UV light was minimised as it causes degradation of nucleic acids. The gel was then washed as follows:

1x10 minutes in 0.25 M HCl
rinse with ddH_2O
1x30 minutes in 0.4 M NaOH:
washes were carried out on a rocking platform.

A 20x20 cm glass plate was placed on top of four plastic supports within a 21x25 cm tray, so that it was raised about 1.5 cm above the bottom of the tray.

The glass plate was covered with two pieces of Whatman 3MM blotting paper (BDH): one 20x20 cm, the other 23x20 cm. The larger piece was folded over, 1.5 cm at each end to overhang the platform and act as a wick.

The blotting paper was carefully wetted with 0.4 M NaOH and 0.4 M NaOH poured into the tray to a level just below the platform.

One piece of Hybond N^+ nylon membrane (Amersham International Ltd., Little Chalfond, Buckinghamshire, UK), four pieces of Whatman 3 MM and five pieces of Quickdraw blotting paper (Sigma) were cut to the same size as the gel.

The gel was placed on the blotting platform and any bubbles excluded by rolling a pipette gently over the gel. The gel was overlayed with the Hybond N^+ membrane, with the top of the membrane level with the bottom of the wells.

Two pieces of 3 MM were soaked in 0.4 M NaOH and placed on top of the membrane. The remaining two pieces of 3 MM and 5 pieces of Quickdraw blotting paper were then placed on top of the blotting platform. A glass plate, clingfilm and a 1 kg weight were placed on top and the apparatus left overnight.

The apparatus was disassembled and the membrane rinsed in 2x SSC for 5 minutes at room temperature. The membrane was air dried on 3 MM for 30 minutes, wrapped in clingfilm and stored at 4°C.

DNA hybridization with Hybond N^+

A nylon mesh was soaked in 2x SSC and a Southern blot membrane placed on top in the 2x SSC solution. Carefully excluding air bubbles, the mesh and membrane were rolled up and inserted into a hybridization bottle (Hybaid Ltd, Teddington, UK). Approximately 20 ml of 2x SSC were added and the bottle rocked to unroll the membrane. The 2x SSC was poured out. A 15 ml volume of hybridization solution, pre-warmed to 65°C and containing denatured salmon sperm DNA at 100 μg/ ml, was added to the bottle, which was then incubated at 65°C on a rotisserie in a hybridization oven (Hybaid Ltd., Teddington, UK) for 1 hour . A radioactive DNA probe, prepared as described above, was heated for 5 minutes at 100°C, before adding to the solution in the hybridization bottle containing the membrane. After hybridization overnight at 65°C, rotating in the hybridization oven, the radioactive contents of the bottle were poured into a 50 ml falcon tube for re-use, allowed to cool to room temperature, then stored at -20°C. The membrane was left in the bottle and washed, rotating in the hybridization oven, as follows:

4x 5 minutes with 100 ml of 2xSSPE, 0.1% SDS solution at room temperature
1x 15 minutes with 100 ml of 1xSSPE, 0.1% SDS solution at 65°C
1x 10 minutes with 100 ml of 0.1xSSPE, 0.1% SDS solution at 65°C

The membrane was blotted briefly between two pieces of 3 MM to remove excess liquid, then wrapped in Saran wrap to prevent drying. The radioactive blot was placed DNA side up in an X-ray cassette and overlayed with Kodak X-OMAT X-ray film (Sigma), and left at -75°C for 72 hours to 1 week.

Southern blots were re-used several times with different radioactive DNA probes, but were stripped between subsequent hybridizations to remove the radioactive DNA from previous probes. The stripping procedure was as follows:

A boiling solution of 0.5% SDS was poured on to membranes and allowed to cool to room temperature with rocking for 1 hour. The membranes were blotted briefly between 3 MM then wrapped in Saran wrap and used for rehybridization.

2.2.9 Cloning DNA in bacterial plasmid vectors

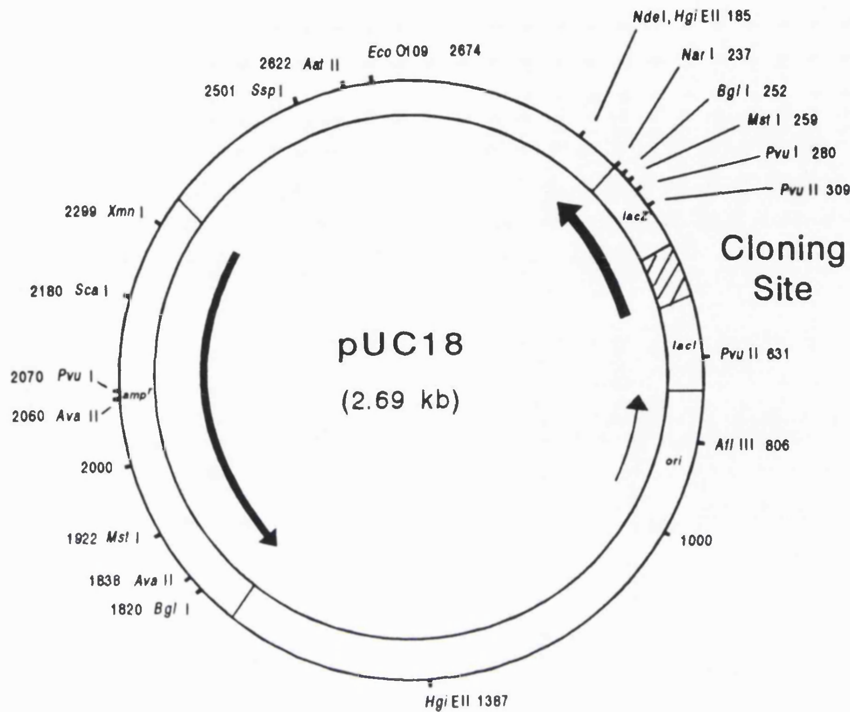
Plasmid vectors

Plasmids were first used as cloning vectors by Cohen *et al* (1973), Hershfield *et al*, (1974) and Covey *et al*, (1976). The early vectors carried the pMB1 or ColE1 replicon, sequences that enabled the double stranded circles of DNA to be replicated in *E.coli*. These early vectors were based on naturally occurring bacterial plasmids, but were modified to carry selectable markers. One problem with these early vectors was that none of them contained more than two restriction sites that could be used for cloning. Further engineering of the plasmids lead to the construction of pBR313, which contained the selectable markers *tet*^{r} and *amp*^{r} (enabling bacteria carrying the plasmid to grow on media containing tetracycline and ampicillin) and numerous unique restriction sites for inserting foreign DNA (Bolivar *et al*, 1977a; 1977b). This plasmid was further refined by removing non-essential plasmid DNA sequences to give the 4.36 kb vector pBR322 (Bolivar *et al*, 1977b). Many of the plasmid vectors used today have been engineered from pBR322 (Balbas *et al*, 1986).

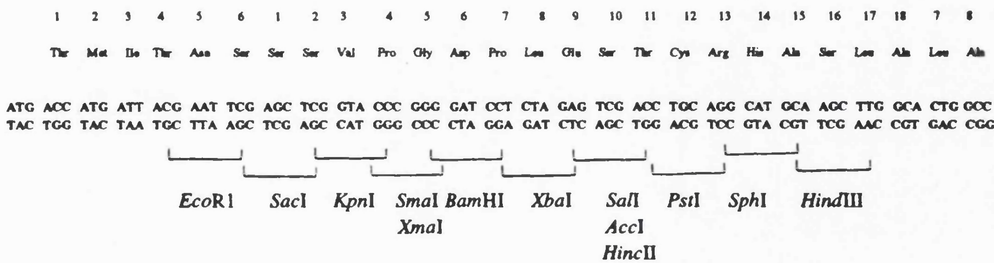
The cloning vector used predominantly in our study was pUC18, which was developed by Messing and co-workers (Vieira & Messing, 1982; Norrander *et al*, 1983; Yanisch-Perron *et al*, 1985), contains sequences from the bacteriophage M13 and is 2.7 kb in size. Bacteria taking up pUC18 during transformation acquire an ampicillin resistance gene encoded by the plasmid. This allows the selection of transformed bacteria on LB-ampicillin agar plates.

In addition pUC18 contains a closely arranged series of cloning sites, termed 'polylinkers', which have been engineered into the plasmid sequence. In most cases the restriction sites are unique: occurring only once in the plasmid sequence. A tandem array of 13 restriction enzyme cleavage sites is found in pUC18: *Eco*RI, *Sac*I, *Kpn*I, *Sma*I, *Xma*I, *Bam*HI, *Xba*I, *Sal*I, *Acc*I, *Hinc*II, *Pst*I, *Sph*I and *Hind*III. These sites provide a variety of targets that can be used singly or in combination to clone DNA fragments generated by cleavage with a large number of restriction enzymes. Fragments inserted at one restriction site can be excised using enzymes that cut at flanking sites. The restriction map of pUC18 is shown in Figure 2.2.

Figure 2.2: Restriction map and polylinker region of the pUC18 plasmid vector.



Cloning site polylinker of pUC18:



The pUC18 plasmid is a circular double stranded DNA molecule, with sequences for an origin of replication, *ori*, a gene conferring ampicillin resistance, amp^r, and the *lacZ'* sequence, which encodes the N-terminal of ß-galactosidase. The above 'polylinker' sequence has been ligated into the beginning of the *lacZ'* region and includes 10 unique restriction enzyme cutting sites. These sites can be used to linearise the plasmid and allow cloning by ligation of DNA inserts with complementary ends. The polylinker adds an additional 18 amino acids to the ß-galactosidase enzyme as shown above. Insertion of DNA into the polylinker causes a shift in reading frame or introduces additional amino acids leading to a non-functional ß-galactosidase enzyme.

In pUC18 the polylinker is positioned within a sequence of DNA derived from the *lac* operon of *E. coli*, which encodes the amino-terminal fragment of ß-galactosidase. The synthesis of this plasmid encoded polypeptide is induced by isopropylthio-ß-D-galactoside (IPTG), and is capable of intra-allelic (α) complemetation with a defective form of ß-galactosidase encoded by the host bacteria (Ullmann *et al*, 1967). In bacterial strains such as JM83 that are transformed with pUC18, the bacteria produce the carboxy-terminal of the ß-galactosidase if induced by IPTG and the plasmid produces the amino-terminal of the enzyme. The two polypeptides form a functional ß-galactosidase enzyme, which cleaves 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal) to give a blue product (Horwitz *et al*, 1964). Colonies derived from pUC18 transformed JM83 that are plated on LB ampicillin, X-gal, IPTG agar plates are thus blue.

The location of the polylinker within the *lac* operon of the plasmid allows for insertional inactivation of the ß-galactosidase enzyme. Insertion of DNA into the N-terminal peptide coding region of the plasmid destroys the production of functional ß-galactosidase, resulting in white colonies on the LB ampicillin, X-gal, IPTG agar plates, as opposed to blue colonies of bacteria containing wild type pUC18.

Initially pUC18 and JM83 host bacteria were used for cloning in the studies described in the following chapters. Whilst transformation and plasmid extraction were efficient, restriction digestion of plasmid DNA often proved difficult. Certain *E.coli*, including the JM series, release high amounts of carbohydrates when lysed in the plasmid preparation procedure detailed later in this section. For this reason the host strain was changed to DH5α (Gibco-BRL, Paisley, Scotland). This *E. coli* strain produces less carbohydrate on alkaline lysis and has the added advantage of not requiring IPTG for induction of ß-galactosidase activity, as the production of the enyzme is constitutive in these cells (Hanahan, 1983).

In order to clone PCR products, the pUC18 vector was digested with *Hinc*II endonuclease, which yields a linear DNA molecule with blunt ends. In certain instances DNA fragments were produced with overhanging single stranded ends, e.g. by digestion with *Eco*RI endonuclease, and these were cloned into pUC18 vector that

had been digested with the same restriction enzyme to give complementary overhanging single stranded ends. Both blunt-ended and cohesive-end ligation techniques are described in the following sections.

Blunt end ligation of DNA into bacterial vectors

Most DNA polymerases used in PCR amplification generate PCR fragments with a single A-residue overhanging the 3' end. For blunt-end ligation this 3' tail was filled in with Klenow polymerase (the large fragment of DNA polymerase I). This procedure was carried out as follows:

The PCR fragment to be cloned was purified from agarose gel and quantified as described in section 2.2.4. The PCR DNA was diluted to an approximate concentration of 200 ng in 20.5 μl ddH_2O. To this DNA was added: 2.5 μl 10x Klenow buffer (0.5M Tris (pH 7.6), 0.1 M $MgCl_2$), 1 μl 2.0 mM dNTPs and 1 μl (6 U) Klenow DNA polymerase (Gibco-BRL, Paisley, Scotland). The solution was gently mixed and incubated at 37°C for 30 minutes. The polymerase was inactivated by heating at 65°C for 10 minutes. The DNA was precipitated and resupended in 5 μl ddH_2O.

The pUC18 vector was prepared for ligation by digesting 2 μg of plasmid in a total volume of 50 μl overnight at 37°C, with 24 U *Hinc*II enzyme (Promega, Southampton, UK), to produce linear plasmid with blunt ends. A 1 μl aliquot of the digested plasmid was electrophoresed on an agarose minigel alongside lambda *Hind*III size markers, to check that digestion was complete. If two bands, corresponding to supercoiled and nicked plasmid, could be seen a further 12 U of *Hinc*II enzyme were added to the remaining plasmid and the digestion continued for a further 5 hours. Complete digestion was indicated by a single band of 2.7 kb, corresponding to the linearized plasmid. The concentration of the plasmid was generally around 40 ng/ μl, determined by electrophoresis on a minigel and comparison with the lambda *Hind*III marker, as described in section 2.2.4.

An equimolar amount of insert and digested plasmid were prepared for ligation. The amount of insert equimolar to 100 ng of linearized plasmid was calculated by the following equation:

$$\text{Amount of insert (ng)} = \frac{\text{Size of insert (kb)}}{\text{Size of plasmid (kb)}} \times 100 \text{ ng}$$

For a 1 kb PCR product and a 2.7 kb plasmid, approximately 40 ng of insert and 100 ng of vector were added together, precipitated and redissolved in a total volume of 6.5 μl of ddH_2O. To this DNA were added: 2 μl of 5x T4 DNA ligase buffer, 0.5 μl 10 mM dATP and 1 μl T4 DNA ligase enzyme, (1 Weiss unit/ μl; Gibco-BRL, Paisley, Soctland). This ligation mixture was incubated at 15°C for 16 hours, during which time the insert and plasmid DNA molecules are joined together by the action of the ligase to form recircularised recombinant plasmid containing the insert.

The association of the ends of the two separate molecules is aided by the low temperature, dATP at 0.5 mM and the presence of polyethylene glycol in the T4 DNA ligase buffer, which causes DNA to condense into aggregates and encourages intermolecular ligation (Pheiffer & Zimmerman, 1983).

Cohesive end ligation

Ligation of DNA with cohesive ends is much more efficient than blunt-end ligation, as hydrogen bonding between complementary single-stranded overhangs at either end of DNA molecules allows them to associate. Cohesive end ligation is particularly useful when high numbers of recombinant plasmids are required or when there is little insert DNA available for cloning. The method used was similar to that for blunt-end ligation, but required less ligase and was carried out as follows:

Insert DNA was digested with an appropriate endonuclease, purified from agarose by electroelution and resuspended in 4.0 μl of ddH_2O. Plasmid DNA was digested with endonuclease to give linear molecules with compatible ends complementary to those of the digested insert DNA. Plasmid DNA (100 ng) and an equimolar amount of

digested insert DNA (around 40 ng for a 1 kb insert) were added to the same tube in a total volume of 6.5 μl ddH_2O. The solution was heated at 50°C for 5 minutes to melt cohesive ends that may have reannealed and the mixture chilled to 0°C. To this DNA solution were added 2 μl 5x T4 DNA ligase buffer, 0.5 μl 10 mM dATP, and 0.1 Weiss units T4 DNA ligase. The ligation mixture was incubated for 16 hours at 15°C.

Transformation of bacteria with plasmid

Mandel & Higa (1970) showed that bacteria treated with ice-cold solutions of $CaCl_2$ and then briefly heat shocked could be transfected with lambda DNA. The same method was subsequently used to transform bacteria with plasmid DNA (Cohen *et al*, 1972). The following procedure was used to transform DH5α *E. coli* with pUC18 plasmid.

A culture of DH5α bacteria was streaked out on an LB-agar plate, which was then incubated overnight at 37°C. A single colony was picked from the plate and transferred into 5 ml of LB broth, which was incubated overnight at 37°C with vigorous shaking.

One ml of the culture was transferred to 50 ml of LB broth in a 500 ml flask and then incubated for 2 hours at 37°C with vigorous shaking. The 50 ml culture was transferred to a sterile, disposable, ice-cold 50 ml Falcon tube, then cooled to 0°C by placing on ice for 10 minutes. The bacterial cells were then pelleted by centrifugation at 3000 rpm for 10 minutes at 4°C and the media decanted into 1% sodium hypochlorite. The tube was inverted for 1 minute to allow the last traces of media to drain away.

The pellet was resuspended in 15 ml of ice-cold 50 mM $CaCl_2$ and placed on ice for 15 minutes. The bacteria were then pelleted by centrifugation for 10 minutes at 3000 rpm at 4°C. The supernatant was decanted off into sodium hypochlorite and the tubes inverted to allow the last traces of fluid to drain away.

The bacterial pellet was resuspended in 4 ml ice-cold 50 mM $CaCl_2$ and left on ice for at least four hours.

A 100 μl aliquot of competent cells was transferred into a prechilled sterile bijou tube. Half the ligation mixture (around 70 ng of DNA) was added to the aliquot of competent cells and mixed gently with the end of the pipette tip. The cells and DNA were placed on ice for 40 minutes. Controls were prepared as follows:

(i) Competent bacteria that were mixed with 10 ng of the standard preparation of supercoiled pUC18 plasmid DNA. This provides a check that the wild type plasmid gives blue colonies on transformation.
(ii) Competent bacteria that were mixed with digested, non-ligated, plasmid DNA, without insert, to check for complete restriction digestion of plasmid.
(iii) Competent bacteria transformed with digested and ligated plasmid to check on efficiency of ligation.

The tubes were transferred to a rack and placed in a circulating water bath at 37°C (this temperature is increased to 42°C if using JM83 *E.coli* cells) and heat-shocked for 90 seconds. The tubes were rapidly transferred to an ice bath, chilled for 2 minutes, then 500 μl of LB broth added to each.

The cultures were incubated for 30 minutes at 37°C in an incubator without agitation, then transferred to a rotary shaker for 30 minutes at 37°C and 180 rpm. This incubation period allowed growth of the bacteria and expression of the plasmid encoded ampicillin resistance protein in transformed cells.

Approximately 150 μl of each transformed culture were plated out on dried, prewarmed LB ampicillin, X-gal, IPTG agar plates, using a sterile glass spreader. The entire culture was either spread onto four plates, or the remains of the culture stored at 4°C and plated the next day, if more recombinant clones were required.

The plates were incubated at 37°C overnight. White colonies containing recombinant plasmid appeared in 12-16 hours, amongst blue colonies containing non-recombinant pUC18. Plasmids were prepared from white colonies as described below.

Plasmid preparation

The following procedure for plasmid preparation is a modification of the methods of Birnboim & Doly (1979) and Ish-Horowicz & Burke (1981).

Cultures were prepared by inoculation of 10 ml LB ampicillin broth with a single colony of plasmid containing bacteria and grown overnight at 37°C with shaking. Glycerol stocks were prepared using 2 ml of each culture and 0.3 ml sterile glycerol, which are mixed well then stored at -70°C.

The cells from the remaining 8 ml of culture were pelleted by spinning at 3000 rpm for 10 minutes at 4°C. After decanting off the supernatant the pellet of cells was resuspended in 0.3 ml of a buffer containing 50 mM Tris (pH 8.0), 10 mM EDTA and RNase A at 100 μg/ ml, making sure there were no clumps of bacteria. The above buffer was stored at 4°C to prevent degradation of RNase.

The cells were then lysed by the addition of 0.3 ml of SDS/ NaOH lysis solution, gentle mixing by inversion 6 times, and incubation at room temperature for 5 minutes. Excessive mixing was avoided as it can shear bacterial DNA and lead to contamination of plasmid preparation. The SDS/ NaOH lysis solution was checked for SDS precipitation, which can occur at low storage temperatures. If necessary the SDS was dissolved by warming at 37°C and the solution allowed to cool to room temperature before using.

The alkaline lysis step was stopped after 5 minutes by the addition of 0.3 ml of 3M potassium acetate (pH 5.5). After addition of potassium acetate, the solution became cloudy and very viscous as potassium dodecyl sulphate (PDS) precipitate, bacterial DNA and proteins came out of solution. To avoid a localized PDS precipitation, the solution was immediately mixed, gently, but thoroughly by inverting 6 times, then incubated on ice for 15 minutes.

The samples were mixed again by inversion then centrifuged at 13,000 rpm for 30 minutes at 4°C and the supernatant removed promptly into a 1.5 ml eppendorfs.

The plasmid was precipitated from the supernatant by adding 2 volumes of absolute ethanol at room temperature and inverting to mix.

The samples were centrifuged at 13,000 rpm for 6 minutes to pellet plasmid DNA. The supernatant was discarded and the tubes inverted on tissue paper to allow the last traces of fluid to drain away from the pellets. The tubes were inverted for around 20 minutes to dry the pellets.

The plasmid DNA pellets were resuspended in 100 μl TE buffer, then extracted with 100 μl of a 1:1 mixture of equilibrated phenol: chloroform, mixing by inversion.

The tubes were spun for 2 minutes at 13,000 rpm and room temperature and the upper, aqueous phase, transferred into 1.5 ml eppendorf tubes.

The plasmid DNA was precipitated by the addition of 2 volumes ethanol, then pelleted by centrifugation for 6 minutes at 13,000 rpm at room temperature. The supernatant was discarded and the tubes inverted on tissue paper to allow the pellets to drain. The plasmid DNA was allowed to dry for 20 minutes, until all traces of ethanol were gone and resuspended in 80 μl TE buffer.

A 3 μl aliquot of each plasmid sample was mixed well with 2 μl of ddH_2O and 1 μl of 10x bromophenol blue/ xylene cyanol loading buffer, followed by electrophoresis on a 0.8% agarose gel in 1x TBE alongside lambda *Hind*III DNA size markers to check the quantity and quality of the plasmid DNA preparations. Plasmids carrying insert DNA are larger and migrate more slowly than the wild type plasmid DNA on electrophoresis. Plasmids with a slower rate of migration on electrophoresis were assumed to carry the desired insert.

Colony hybridization

A certain proportion of false white colonies were produced during the ligation and transformation process, which had no insert, but small deletions in the vector within the ß-galactosidase gene. The deletion of DNA from the α-peptide encoding region can give the same result as if there were an insertion, i.e. non-functional ß-galactosidase activity. The proportion of false white colonies was determined by carrying out twelve plasmid preparations, as described above, and looking for plasmid samples that migrated at the same rate as wild type pUC18 on 0.8% agarose minigel electrophoresis. If there was a high proportion of false white colonies, the bacteria carrying recombinant plasmid DNA were detected by colony hybridization. The colony hybridization procedure was developed by Grunstein and Hogness (1975) and is detailed below.

White colonies of bacteria were picked with sterile cocktail sticks and gridded out 200 to a 90 mm LB ampicillin agar plate. The plates were left overnight at 37°C.

A nitrocellulose membrane (Anderman, Kingston, Surrey, UK) was cut into an 85 mm circle. Matching orientation lines were drawn on the side of the bacterial agar plate and on the nitrocellulose membrane to allow the precise positioning at the end of the procedure.

The nitrocellulose membrane was pre-moistened on sterile LB ampicillin agar plates and layed on top of the colonies for 1 minute, with the lines on the membrane matching lines on the agar plate, then removed and placed bacteria face up on a fresh LB ampicillin plate.

The bacteria thus transferred were grown overnight at 37°C until large colonies were present on the nitrocellulose membrane. The original plate was incubated for 5-7 hours at 37°C to regenerate the colonies, then sealed with parafilm and stored at 4°C.

The colonies on the membrane were placed face up on Whatman 3MM paper soaked

in 10% SDS for 3 minutes to help maintain the integrity of the colonies.

The bacteria were lysed on 3MM paper soaked in S1 solution (see section 2.1.4) for 5 minutes, again with the bacteria face up. The membrane was then transfered to 3MM paper soaked in S2 for 5 minutes, then washed briefly through 2x SSC.

The membranes were allowed to air dry for 30 minutes, then were placed in an oven at 80°C for 2 hours to bind the DNA.

The blots were incubated at 55°C in 100 ml of a proteinase K buffer containing 400μg of proteinase K (Promega, Southampton, UK), for 1 hour.

Hybridization of the membrane overnight with a radioactive cDNA probe complementary to the inserted DNA, followed by washing, and exposure of the membrane to Fuji X-100 X-ray film (GRI, Dunmow, Essex, UK) for 1-12 hours revealed which colonies contained recombinant plasmid, see section 2.2.8 for hybridization and washing details.

The orientation lines on the membrane were drawn on the autoradioagraphy film and then matched up with the lines on the bacterial plate to show which colonies had insert DNA.

Colonies believed to contain insert were picked from the original plates, grown overnight in 10 ml of LB ampicillin broth and plasmid DNA extracted, using the method given above. A glycerol stock was prepared for each culture, with glycerol at 15%, and stored frozen at -20°C.

2.2.10 Dideoxy chain termination DNA sequencing

Sequencing double-stranded plasmid DNA

Sequencing of double-stranded plasmid DNA was carried out using a modified version of the dideoxy chain termination method (Sanger *et al*, 1977) and the Sequenase version 2.0 kit (United States Biochemicals, Cambridge Bioscience, Cambridge, UK) by the following procedure:

Approximately 2.5 μg of plasmid DNA (20-40 μl of plasmid preparation from an 8 ml culture) was made single stranded in 0.2 M NaOH, by the addition of 5 μl of 1M NaOH to 20 μl of plasmid solution and incubation at 37°C for 30 minutes.

One tenth the volume of 3M Na acetate (pH 5.2) was then added to neutralize the NaOH, 2.5 times the volume of ethanol added and the DNA precipitated at -75°C for 15 minutes.

This single stranded template was spun down at 13,000 rpm for 15 minutes, the supernatant poured off, the tube inverted on tissue paper and the pellet allowed to air dry for 10 minutes.

The precipitated single stranded plasmid DNA was resuspended in 7 μl of ddH_2O, 2 μl of manganese buffer added, along with 1 μl of 0.5 mM M13 'universal' primer and the primer annealed by heating at 37°C for 15 minutes.

The annealing mixture was added to 5.5μl of ^{35}S-dATP labelling master mix solution, consisting of:

1 μl 100 mM dithiothreitol
2 μl 1.5 μM dNTPs
2 μl of diluted Sequenase version 2.0 DNA polymerase enzyme
0.5 μl ^{35}S-dATP 10 μCi/ μl

The Sequenase version 2.0 DNA polymerase enzyme was diluted 1:7 with the enzyme dilution buffer provided, just prior to addition of annealing mix. The combined

annealing mix and labelling mix were incubated at room temperature for 2 minutes.

Aliquots of 3.5 μl of solution were then added to each of four tubes containing 2.5 μl of a single dideoxynucleotide solution and 80 μM dNTPs: one tube contains only dideoxy-GTP, one dideoxy-ATP, one dideoxy-TTP and the last dideoxy-CTP. The four tubes, appropriately labelled G, A, T, and C were then incubated at 37°C for 5 minutes.

The dideoxy chain termination reaction was stopped by the addition of 4 μl of a solution containing 95% formamide and the tracking dyes bromophenol blue and xylene cyanol.

Sequencing electrophoresis conditions

Gels were cast between 20x30 cm glass plates, one of which was silanised with 5% dimethyldichlorosilane in chloroform. Two 20x30 cm spacers (0.4 mm in thickness) were placed between the plates, which were then taped up with electrophoresis tape The gels were 6% polyacrylamide containing 7 M urea and were made up in 1x TBE for sequencing buffer (see section 2.1.6).

Polyacrylamide gel electrophoresis (PAGE) was carried out in 1x TBE for sequencing buffer. The gels were preheated for 15 minutes at 48 W. Half of the sequencing samples (4.5 μl) were loaded then the electrophoresis was continued at 40 W for 15 minutes, 33 W for 45 minutes and 28 W for 30 minutes. This maintained the gel at around 50°C. The second half of the sample was then loaded and the gel run at 28 W for a further 1.5 hours, to give a 1.5 and a 3 hour run.

The 1.5 hour run allowed the generation of a sequence that included the polylinker region of the pUC18, the point at which the insert DNA was ligated into the plasmid and the first 100 nucleotides of the insert. The 3 hour run allowed greater separation of larger pieces of DNA, so that another 200-300 nucleotides could be read off.

Once electrophoresis was complete the gels were removed from the tank, the tape removed from the glass plates and the plates prised apart. The gels were fixed in 10% methanol/ 10% acetic acid in ddH_2O for 15 minutes, then dried down at 80°C on a vacuum gel dryer and exposed to Fuji X-100 film (GRI, Dunmow, Essex, UK) for 24-72 hours.

Sequencing PCR products

Double stranded DNA produced by PCR amplification was made single stranded by boiling and prevented from reannealing using non-ionic detergents, such as 10% DMSO (Winship, 1989) or a mixture of 0.5% NP40 and 0.5% Tween 20 (Keohavong & Thilly, 1989; Bachmann *et al*, 1990). The PCR DNA was boiled for 2 minutes in the presence of detergent, manganese buffer and primer, snap cooled on ice and the labelling master mix added as above. Incubation with labelling mix was carried out at room temperature for 5 minutes. Aliquots of 3.5 μl were then added to each of the four dideoxy termination tubes, which also contained detergent at the percentages given above. The termination reaction was allowed to proceed for 5 minutes before stopping with 95% formamide/ bromophenol blue solution. When using the method of Bachmann *et al* (1990), 2 μl of a solution containing 0.25 mM dNTP, 50 mM NaCl, 0.5% NP40 and 0.5% Tween 20 was added after the first 5 minutes of dideoxy termination and the reaction allowed to continue for another 5 minutes before stopping with 95% formamide solution. The products of the sequencing reaction were analyzed as described in the section on sequencing electrophoresis conditions.

Taq cycle sequencing

Sequencing was also carried out by *Taq* cycle sequencing, using the method of Embley (1991), and the TAQuence kit (United States Biochemicals, Cambridge, UK). Two picomoles of primer was end-labelled using 6 units of T4 polynucleotide kinase (United States Biochemicals, Cambridge, UK) and 1 μl of gamma ^{32}P-dATP (3000 Ci/ mmol, 10 μCi/ μl; NEN, Stevenage, UK) in 1x T4 polynucleotide kinase buffer in a volume of 10 μl, according to manufacturers instructions. Each

sequencing reaction contained 3 μl template DNA (50-200 ng), 9.5 μl ddH_2O, 1 μl labelled primer (0.2 pmol), 2 μl Taquence reaction buffer, and 2 μl of diluted (1 unit/ μl) *Taq* polymerase. This mixture was spun down and 4 μl added to each of four tubes containing 4 μl of termination mixture (either ddATP, ddCTP, ddGTP, ddTTP). Each tube was overlayed with 30 μl of light mineral oil. The thermal programme for sequencing consisted of 10 cycles of denaturation at 95°C for 40 seconds, 30 seconds of annealing at 58°C and 2 minutes extension at 72°C, with the single gamma ^{32}P-dATP end-labelled primer. The sequencing reactions were stopped by adding 4 μl of TAQuence stop mix to each tube. The products of the sequencing reactions were analyzed (6 μl) on standard sequencing gels, as detailed above.

3- Gene library screening for mutations in a case of 17α-hydroxylase deficiency

3.1 Clinical details

This female patient presented at age 29 years with non-responsive hypertension and lack of secondary sexual characterisitics. A diagnosis of 17α-hydroxylase deficiency was made on the basis of low plasma cortisol concentrations (107 nmol/l at 0800h, normal range 165-715 nmol/l) and raised plasma levels of corticosterone (972 nmol/l, normal range 2-23 nmol/l), deoxycorticosterone (4906 pmol/l, normal range 90-482 pmol/l) and ACTH (34 pmol/l, normal range 4-13 pmol/l) (Fraser *et al*, 1987). Analysis of chromosomes revealed a normal karyotype, 46XX. The patient was tall (190 cm) and had no breast development or pubic hair. The vulva was infantile, the labia were small, but the clitoris was normal. The vagina had a rugose appearance and the introitus admitted two fingers. No uterus was palpable, but on laparoscopy the uterus was seen as a small fold. Streak gonads were visible.

3.2 Experimental work

3.2.1 Southern blot analysis

Southern blot analysis was carried out to ascertain if there were large deletions or insertions within the *CYP17* gene of the subject. Genomic DNA from peripheral leucocytes was isolated, 10 μg digested with *Eco*RI restriction endonuclease and after Southern blotting, hybridized with the ^{32}P-labelled *Bam*HI insert from the CYP17 cDNA, pCD17αH, kindly supplied by Professor M. Waterman of Vanderbilt University, Nashville (see section 2.2.8 for details of method).

Two DNA fragments of 5.7 kb and 6.9 kb were observed (Figure 3.2a), which are the expected size for *Eco*RI digested *CYP17* (Picardo-Leonard & Miller, 1987). This result excluded any large scale deletion or insertion as the cause of disease.

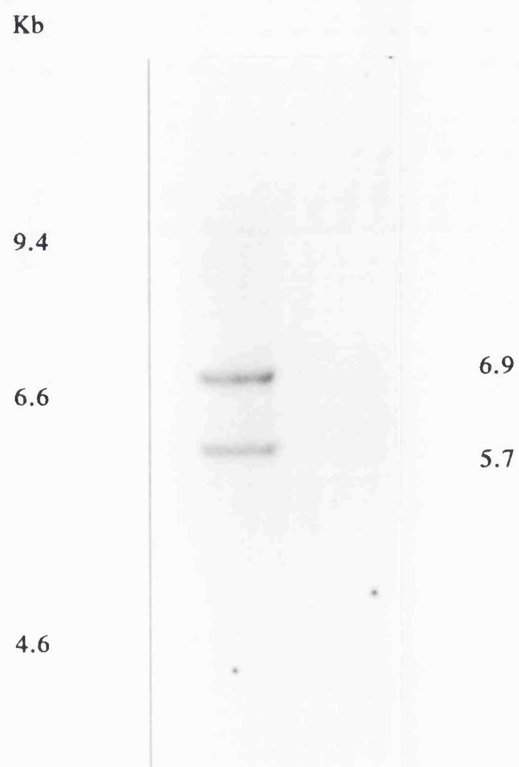
3.2.2 Strategy for locating mutations within *CYP17*

The lack of any large alteration within the *CYP17* gene suggested that smaller mutations, possibly involving single base changes, were responsible for the loss of enzyme function. To locate these mutations it was necessary to isolate the *CYP17* gene from genomic DNA. A gene library was constructed by inserting *Eco*RI digested, size-selected, genomic DNA from the subject into bacteriophage lambda gt10. This library, containing numerous gene fragments, was then screened to isolate bacteriophage containing either the 5.7 kb or 6.9 kb *Eco*RI fragment of *CYP17*. The procedure is outlined in Figure 3.2b and described in detail in the following sections.

Bacteriophage lamda gt10

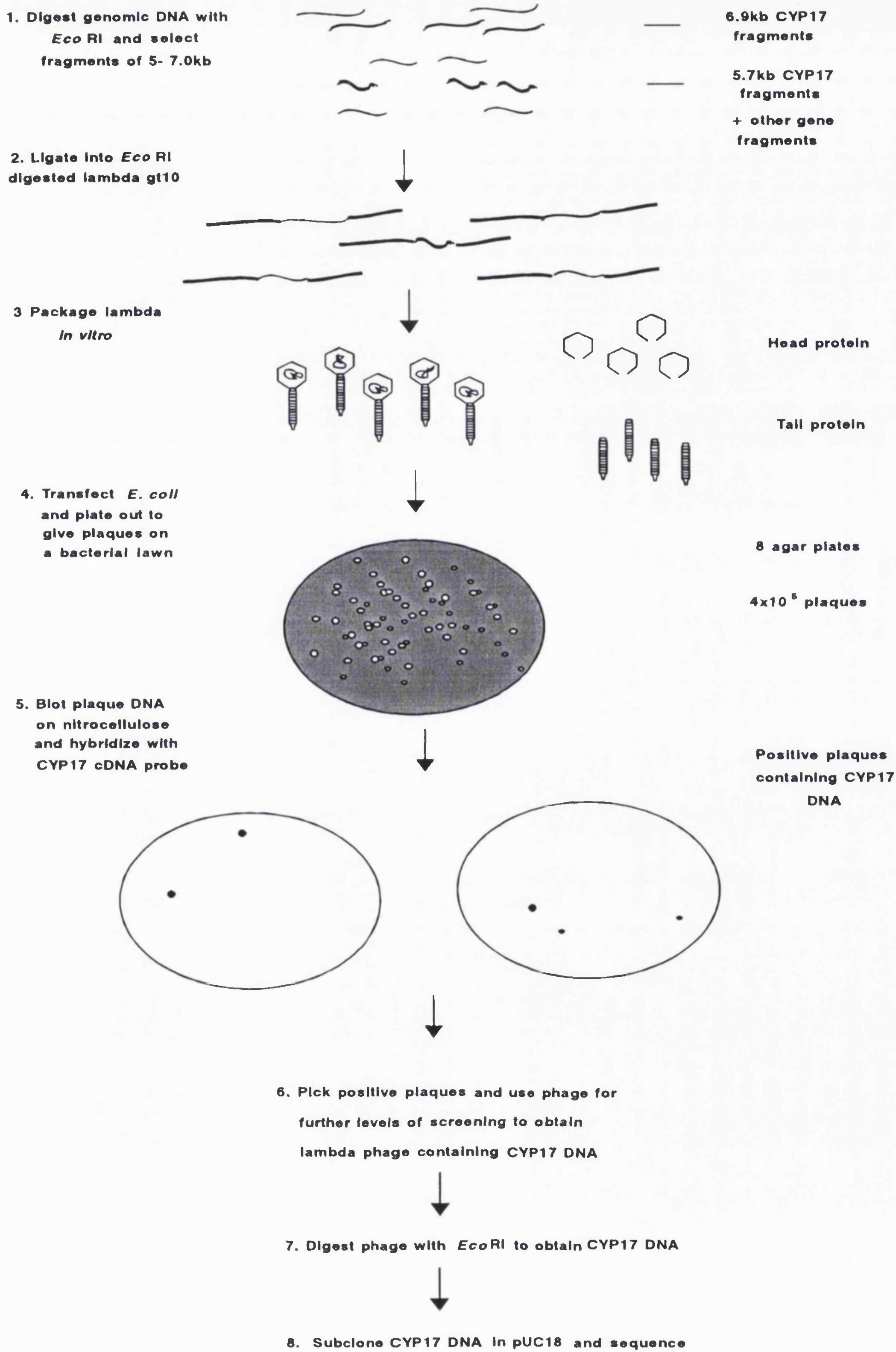
Wild type bacteriophage lambda is a 50 kb double-stranded DNA virus that infects *E. coli*. Lambda gt10 is derived from the wild type bacteriophage, but has been designed for cloning small DNA fragments, up to 6 kb, into a unique *Eco*RI site within the gene for the *c*I protein. Cloning efficiency is high and recombinants are selected on an *E.coli* strain, such as C600 *hfl*A, carrying the *hfl*A150 mutation, which gives high frequency of lysogeny. Non-recombinant bacteriophage integrates into the host DNA and the bacteria remain in a state of lysogeny due to an active *c*I protein derived from the phage. Recombinant bacteriophage has insertional inactivation of the *c*I gene and therefore can grow normally, enter a lytic cycle and form plaques that contain phage DNA with insert (Huynh *et al*, 1985). Bacteriophage DNA from plaques can be isolated by laying nitrocellulose membranes on top of lawns of infected bacteria. The blots produced are hybridized with a radioactive cDNA probe complementary to the gene of interest, to detect which plaques contain bacteriophage with the desired inserted DNA.

Figure 3.2a: Southern blot analysis of genomic DNA following *Eco*RI digestion



In the Southern blot above, 10 μg of *Eco*RI digested genomic DNA from the 17α-hydroxylase deficient patient was loaded alongside lambda *Hind*III size markers. The expected band sizes of 5.7 and 6.9kb. were observed on hydridization as shown.

3.2b: Strategy for screening a lambda gt10 gene library for *CYP17* gene fragments.



3.2.3 *Eco*RI digestion of genomic DNA and size selection by electrophoresis

Approximately 100 μg of genomic DNA, obtained from peripheral leucocytes, was digested overnight at 37°C with the restriction enzyme *EcoRI*, in a 300 μl reaction volume. A 1 μl aliquot of the digestion was taken and electrophoresed on a 0.8% agarose minigel to check for complete digestion. A 20x20 cm 0.8% agarose gel was prepared in 1x TAE, using a comb that produced two small wells of 35 μl capacity on either side of a single large well capable of holding 300 μl. The digested DNA was mixed with 33 μl of 10x bromophenol blue/ xylene cyanol loading buffer. A 30 μl aliquot was loaded into the small well, the rest of the digested genomic DNA was loaded into the large well, and lambda *Hind*III DNA size markers loaded into the other small well of the 20x20 cm agarose gel. Electrophoresis was carried out at 46 V and 40 mA overnight in 1x TAE.

The gel was removed from the tank and viewed briefly on a UV transilluminator. The region of the gel containing the aliquot of around 10 μg of digested genomic DNA was carefully cut away and the gel Southern blotted. The migration distances of the lambda markers were measured and the area of gel between the 4 kb and 9 kb size markers excised, covered with cling film and stored at 4°C. The Southern blotting of the small section of gel determined which region of the gel contained the 5.7 and 6.9 kb fragments of the *CYP17* gene.

When the exact position of the 5.7 and 6.9 kb bands had been determined the corresponding regions on the stored gel were excised as thin bands of agarose, which were cut into small pieces of around 5x3 mm. The DNA was purified from the agarose by electroelution as described in section 2.2.7. Approximately 400 ng of DNA was recovered.

3.2.4 Ligation of lambda bacteriophage arms to *Eco*RI digested DNA

Four ligations were set up, each with 500 ng of the *Eco*RI digested lambda bacteriophage right and left arms (supplied in the Gigapack Gold kit; Stratagene, Cambridge, UK), but containing 20, 50, 100 and 200 ng amounts of the purified, size selected genomic DNA, representing molar ratios of insert: vector DNA of 1:3, 2:1 and 4:1. A control ligation containing only the right and left arms of lambda was also prepared. The left arm of *Eco*RI digested lambda is 32.7 kb in length and the right arm 10.6 kb. Ligations were carried out in 10 μl volumes with 0.2 Weiss Units of T4 DNA ligase, overnight at 16°C as described in section 2.2.9.

Following ligation 2 μl aliquots of each ligation were electrophoresed on a 0.5% agarose gel in 1X TBE, alongside wild type lambda to check the efficiency of ligation. Unligated arms and insert were observed in all reactions, but in the ligations containing 100 and 200 ng of insert a small proportion of ligated product similar in size to wild type lambda bacteriophage was observed. The control ligation, containing just left and right arms showed no ligated product, which was expected as the 5' termini of the supplied vector DNA had been dephosphorylated by treatment with alkaline phosphatase.

3.2.5 *In vitro* packaging of lambda bacteriophage.

Packaging of bacteriophage lambda DNA *in vitro* was first developed by Becker & Gold (1975) and utilizes extracts from bacteria infected with bacteriophage lambda mutants. The Gigapack Gold kit contains solutions of packaging proteins for gt10. One extract, obtained by sonicating a strain of bacteria such as BHB2690, after infection by a certain lambda mutant, contains preheads: proteins that will encapsulate the lambda DNA. The other extract contains proteins required to package the DNA into the preheads and is obtained by freeze/ thawing a bacterial strain such as BHB2688, after infection by a lambda mutant that overexpresses packaging proteins (Sambrook *et al*, 1989c). Packaging is carried out *in vitro* by mixing the two extracts and then adding lambda DNA (Scalenghe *et al*, 1981), to form infective viral particles.

In vitro packaging of recombinant lambda gt10

The two ligations that gave observable products were pooled and used for *in vitro* packaging using the Gigapack Gold kit. The sonic extract and freeze/ thaw extracts were removed from -70°C and placed on dry ice. The freeze/ thaw extract was quickly thawed between fingers until just beginning to thaw.

A 5 μl aliquot of the ligation mixture, approximately 300 ng of DNA, was added to the freeze thaw extract and placed on ice. A 15 μl aliquot of sonic extract was added, stirred well to mix and incubated for 2 hours at 22°C.

After incubation, 500 μl SM buffer, then 20 μl chloroform were added to the packaging mixture and gently mixed with a pipette. Centrifugation for 2 minutes at 13,000 rpm at room temperature was carried out to sediment debris. The supernatant was stored at 4°C.

3.2.6 Preparation of competent C600 *hfl*A cells

A glycerol stock of C600*hfl*A cells was streaked out onto an NZY agar plate and incubated overnight at 37°C.

A single colony was used to inoculate 50 ml TB medium supplemented with 10 mM $MgSO_4$ and 0.2% maltose. The maltose is required to induce the bacterial production of *lam* maltose receptors, to which lambda phage binds prior to infection. The culture was incubated at 37°C with vigorous shaking for 4 hours. Absorbance readings were taken at 600 nm and incubation discontinued when absorbance reached 0.80.

The bacteria were pelleted by centrifugation at 3000 rpm for 10 minutes at 4°C. The supernatant was poured off into 1% sodium hypochlorite solution and the pellet resuspended in 25 ml sterile 10 mM $MgSO_4$. The cells were then diluted to an OD_{600} of 0.5, by the addition of 35 ml of sterile 10 mM $MgSO_4$. Glycerol was added to a concentration of 15% and the culture mixed well, before dividing into 0.6 ml aliquots. The competent glycerol stocks were stored at -70°C.

3.2.7 Determination of infectivity of packaged lambda gt10

Three 0.6 ml aliquots of competent C600 *hfl*A cells were thawed on ice and transferred to three 50 ml falcon tubes. To one aliquot of cells was added 1 μl of packaged phage solution, to another 3 μl, and to the third 5 μl. The tubes were then incubated for 20 minutes at 37°C, to allow adsorption of phage.

A volume of 60 ml top agar, containing $MgSO_4$ at 10 mM, was melted by heating in a boiling water bath and cooled to 48°C. 20 ml of molten top agar was added to each of the above tubes of competent cells, and mixed by inversion. The agar was poured onto three dry, prewarmed, 150 mm diameter LB $MgSO_4$ agar plates and allowed to set before incubating at 37°C overnight.

Following incubation a lawn of bacteria with clear areas, or plaques, caused by lysis of bacteria by recombinant lambda phage, was observed. No bacterial growth was observed from the cultures transfected with 3 μl and 5 μl of packaged phage, presumably due to excessive lysis. The culture transfected with 1 μl of packaged phage gave almost confluent plaques, with a few small areas of bacterial growth. The number of lambda plaques giving confluent lysis on a 150 mm plate is stated as 50,000 (Sambrook *et al*, 1989d). From this value the infectivity of the packaged lambda gt10 mixture was approximately 50,000 plaque forming units (pfu)/ μl.

3.2.8 DNA hybridization screening of a lambda gene library

The plaques were screened by transfer of bacteriophage to a nitrocellulose membrane and hybridization with a ^{32}P-dCTP labelled probe complementary to *CYP17* using the method of Benton & Davis (1977). The detailed procedure is given below.

A 600 μl aliquot of C600 *hfl*A competent cells was thawed on ice. Packaged phage solution containing 2x 10^5 pfu, 3 μl in volume, was added to the cells, which were incubated at 37°C for 20 minutes to allow adsorption of the phage.

The transfected cells were added to 80 ml molten LB $MgSO_4$ top agar at 48°C. This was inverted to mix and 20 ml poured onto each of four 150 mm diameter LB $MgSO_4$ agar plates.

The plates were incubated overnight at 37°C to generate near confluent plaques in a lawn of bacteria. The plates were stored at 4°C for at least 1 hour to harden the top agar and prevent it adhering to the nitrocellulose membrane in the following procedure.

Nitrocellulose membrane was cut into four 150 mm diameter circles and orientation marks carefully drawn on to the membrane with pencil. The nitrocellulose membranes were pre-moistened on sterile LB $MgSO_4$ agar plates, then layed on top of the plaques on the four chilled agar plates for 2 minutes at room temperature, to adsorb phage from the plaques. Orientation marks matching those on the nitrocellulose membranes were drawn on the bottom of the agar plate.

The membranes were peeled off and placed DNA face up on blotting paper soaked in S1 solution for 7 minutes to denature phage proteins and DNA. The agar plates were sealed and stored at 4°C.

The membranes were neutralized by placing on blotting paper soaked in S2 for 2x 3 minutes, moving the blots to different parts of the S2 soaked paper.

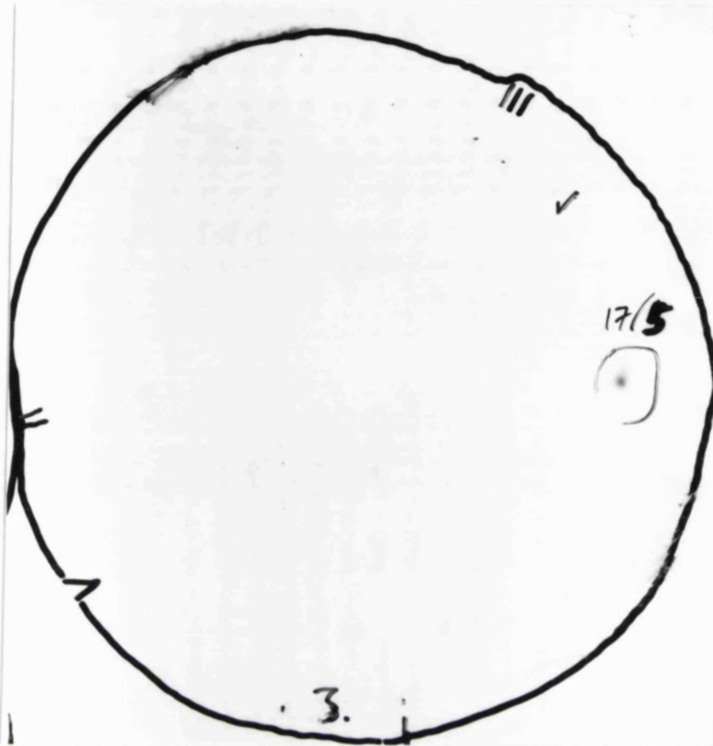
The blots were air dried for 30 minutes, then baked at 80°C for 2 hours to bind lambda DNA to the membrane. The membranes were then incubated at 55°C for 1 hour in 100 ml of proteinase K buffer, containing 400 μg of proteinase K.

The blots containing phage DNA were hybridized overnight with ^{32}P-labelled *CYP17* probe at 65°C, then washed, as descibed in section 2.2.8, to remove excess probe.

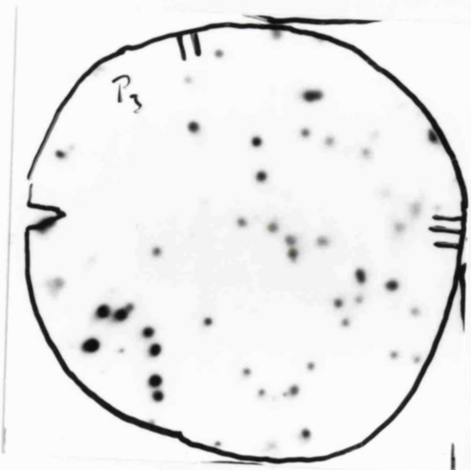
Exposure to X-ray film for up to 72 hours revealed which plaques contained *CYP17* DNA. The orientation marks on the nitrocellulose membranes were marked on the autoradiographs and aligned with the stored agar plates to locate the plaques containing *CYP17* DNA. Examples of autoradiographs showing positive plaques are shown in Figure 3.2c.

The area of top agar believed to contain *CYP17* DNA was removed with a pipette tip and placed in 1 ml of SM buffer in a 1.5 ml eppendorf. The tube was inverted several times to mix, then 100 μl removed into 900 μl of SM to give a 10^{-1} dilution. Serial dilutions in SM buffer were prepared to give diluents of 10^{-4}, 10^{-5} and 10^{-6}. 30 μl of chloroform was added to the first tube containing the picked agar, to prevent bacterial growth and the phage diluents stored at 4°C.

Figure 3.2c: Autoradiographs of primary and secondary screen blots showing positive plaques hybridizing to radioactive ^{32}P-dCTP labelled *CYP17* cDNA.



Example of one of five positive plaques generated by lambda gt10 containing *CYP17* DNA. Autoradiograph is from a nitrocellulose blot of a primary screen agar plate.



Autoradiograph of nitrocellulose blot of a secondary screen agar plate, showing increased proportion of *CYP17* containing plaques.

The pfu values of the phage diluents were determined as before. The 10^{-5} dilution was found to have a value of approximately 100 pfu/ µl. 10 µl of this 10^{-5} diluent was used for a secondary screen, using the same procedure as outlined above. Two further rounds of screening were necessary to obtain agar plates containing plaques with only *CYP17* insert. Five clones containing *CYP17* DNA were isolated from a total of 5x 10^5 plaques screened.

3.2.9 Isolation of recombinant lambda gt10-*CYP17* DNA.

A single plaque, containing lambda with only *CYP17* insert DNA, was picked from a quaternary screen agar plate and placed into 100 µl of SM buffer. The agar was broken up with a pipette tip, the solution gently mixed with a pipette and a 10^{-2} diluent prepared in SM buffer. This solution contained approximately 10^4 pfu/ µl.

Two LB $MgSO_4$, 0.8% agarose plates, 150 mm diameter, with near confluent *CYP17*-containing plaques were prepared using 200 µl of competent C600*hfl*A cells and 10 µl of the 10^{-2} phage diluent, as described above. The lambda obtained was to be digested with *Eco*RI to release the *CYP17* insert, and as restriction endonuclease activity is inhibited by carbohydrate polymers found in agar, the plates used to isolate lambda were prepared using agarose at 0.8% in place of bactoagar.

The plates were each washed with 5 ml SM buffer for 1 hour on a rocking platform at 4°C. The resultant 10 ml solution was taken off, 20 µl of chloroform added and the solution vortexed for 20 seconds, before spinning at 3000 rpm for 10 minutes at 4°C to remove bacterial debris and insoluble agarose contaminants. The supernatant was collected into a sterile 15 ml tube.

Purification of the lambda phage was carried out according to the protocol of the Qiagen lambda kit (Hybaid, Teddington, UK), using the solutions provided. 30 µl of a solution containing RNase A at 20 mg/ ml, DNase I at 6 mg/ ml in 10 mM EDTA, 100 mM Tris, 300 mM NaCl, 0.2 mg/ ml BSA, pH 7.5, was added to the supernatant. The tube was incubated at 37°C for 30 minutes. This digests nucleic acids liberated from lysed bacteria. Without digestion a significant number of

bacteriophage particles become entrapped in the viscous solution of nucleic acids and the bacteriophage yield is reduced.

The bacteriophage particles were precipitated by the addition of 5.0 ml of a 30 % polyethylene glycol/ 3 M NaCl solution (final concentration 10% PEG, 1M NaCl), with incubation on ice for 1 hour.

The bacteriophage was pelleted by spinning at 13,000 rpm for 15 minutes at 4°C. The supernatant was removed, the tube respun at 13,000 rpm for 2 minutes at 4°C and the last remnants of supernatant removed.

The pellet was resuspended in 1 ml of a solution containing 100 mM Tris (pH 7.5), 100 mM NaCl and 25 mM EDTA then divided into two 1.5 ml eppendorfs. 0.5 ml of 4% SDS solution was added to each tube. The solutions were mixed gently and the tubes incubated at 70°C for 10 minutes to dissociate proteins from the lambda DNA. The tubes were placed on ice and allowed to cool for 5 minutes. The SDS and protein was precipitated by the addition of 0.5 ml of 2.55 M potassium acetate (pH 4.8) to each tube, which forms a white precipitate of potassium dodecyl sulphate. The tubes were mixed by inversion, then centrifuged at 13,000 rpm to pellet the precipitate. DNA in the supernatant was precipitated by the addition of 2 volumes of absolute ethanol, followed by centrifugation at 13,000 rpm for 15 minutes at 4°C. The supernatant was poured off and the tubes inverted on tissue paper for 20 minutes to dry. The pellets were resuspended in 43 μl of sterile ddH_2O.

3.2.10 Isolation of *CYP17* DNA and subcloning into pUC18 plasmid vector

The bacteriophage DNA isolated above was digested overnight at 37°C with *Eco*RI restriction endonuclease in a total volume of 50 μl with 24 U of *Eco*RI. The 5.7 kb *Eco*RI fragment of *CYP17* was separated from the vector by electrophoresis, the DNA excised from the minigel and purified by spinning through glass wool. Approximately 20 ng of purified *CYP17* insert was ligated with 50 ng of *Eco*RI digested pUC18 vector at 16°C for 6 hours, using 0.2 U T4 DNA ligase in a total volume of 10 μl. Ligated DNA was transformed into JM83 *E.coli* made competent

in ice cold 50mM $CaCl_2$, then plated out on LB ampicillin, IPTG, X-gal agar plates and incubated overnight at 37°C.

White bacterial colonies were picked with sterile cocktail sticks and gridded out, 200 to a single 90 mm diameter LB ampicillin plate. Colony hybridization using the pCD17αH cDNA probe was then carried out to ascertain which colonies contained *CYP17* DNA (Grunstein and Hogness, 1975). Positive colonies were identified, picked from the original plate, each grown overnight in 10 ml of LB ampicillin broth and plasmid DNA extracted using standard procedure given in section 2.2.9. Electrophoresis on a 0.8% agarose gel alongside wild type pUC18 confirmed that large inserts were present in the extracted plasmids, which on digestion with *Eco*RI and *Pst*I enzymes yielded fragments of 2.6 and 1.7 kb in length, in agreement with the restriction map of the 5' end of *CYP17* (Picardo-Leonard & Miller, 1987). These fragments were subcloned into *Pst*I digested pUC18 and then sequenced as described in section 2.2.10.

Each of the five lambda clones, identified in section 3.2.8, was processed as described above and found to contain the 5.7 kb *Eco*RI fragment of *CYP17*, which included the 5' promoter region and first three exons of *CYP17*.

3.2.11 Results of sequencing *CYP17* DNA obtained from gene library

Exon 3 contained a single point mutation: a G to T transversion, which introduced a stop codon (TAG) in place of glutamate (GAG) at codon 194 (Rumsby *et al*, 1992a). Two silent mutations were also found in exon 1, CAC in place of CAT encoding histidine 46 and TCG in place of TCT encoding serine 65, and have been previously reported (Kagimoto *et al*, 1988). These silent changes were not reported by Picardo-Leonard & Miller (1987) and Bradshaw *et al* (1987), but may represent polymorphisms. Exon 2 was shown to be normal as compared to the published sequences listed above. Four of the isolated clones carried the mutation in exon 3, but the fifth had the normal sequence at this point, indicating that the patient carried two different mutant alleles.

3.2.12 PCR amplification of *CYP17* DNA

To confirm the presence of a premature stop codon at position 194, exons 3 and 4 of *CYP17* were amplified from genomic DNA using PCR. Four 25 μl reactions were prepared, each containing approximately 250 ng of genomic DNA, 10 mM Tris (pH 8.0), 50 mM KCl, 200 μM each of dGTP, dATP, dTTP and dCTP, 1.5 mM $MgCl_2$, 0.25 μM of each of the two primers and 0.5 U *Taq* polymerase, (AmpliTaq; Amersham International, Little Chalfond, Buckinghamshire, UK). The primers used are detailed in Appendix 1. Amplifications were carried out on a Hybaid thermal cycler programmed as follows: denaturation 98°C for 5 minutes, addition of *Taq* polymerase at 60°C, followed by 35 cycles of 1 minute at 72°C, 0.01 minutes at 93°C and 0.01 minutes at 60°C. The double stranded product was purified from agarose and used for *Taq* cycle sequencing using the Taquence kit as detailed in section 2.2.10. The product of a separate PCR reaction, containing exons 3 and 4, was cloned into pUC18 and sequenced.

3.2.13 Results of sequencing PCR amplified DNA

The *Taq* cycle sequencing showed both normal and mutant sequence at codon 194, confirming that the mutation was present on only one allele (Figure 3.2d). A second premature stop codon, CGA→TGA, Arg→Stop at codon 239, was also detected which again appeared to be on only one allele. Two clones containing the PCR amplified exon 3 and 4 fragment were sequenced and showed a normal sequence at position 194, but had the mutation in exon 4, which demonstrated that the two premature stop codons were present on separate alleles (Figure 3.2e).

Figure 3.2d: *Taq* cycle sequencing revealed both normal sequence, GAG, and mutant sequence, TAG, at codon 194, showing the E194X mutation to be present on only one allele.

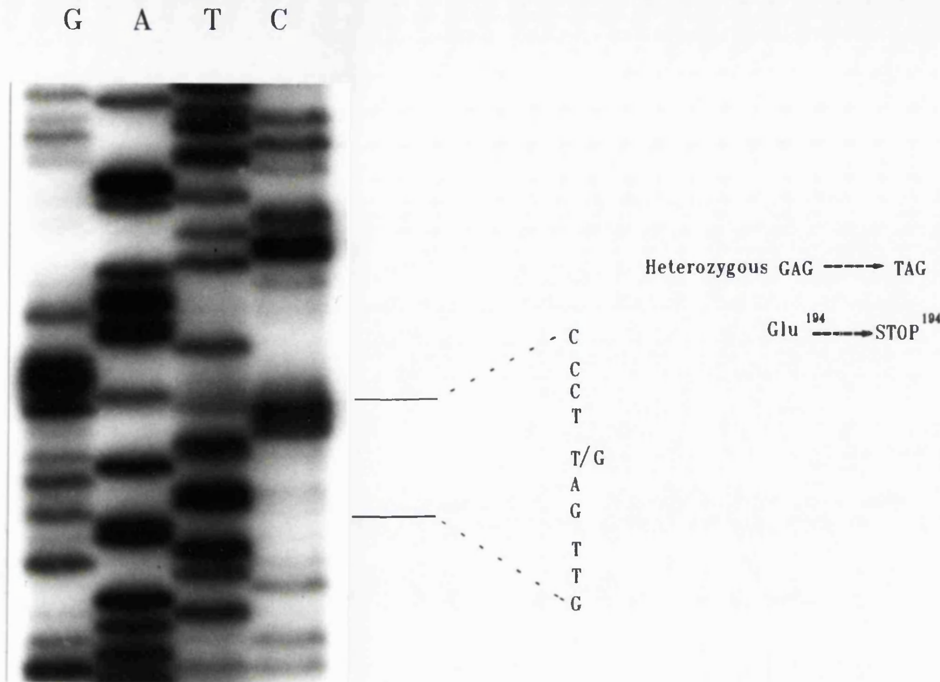
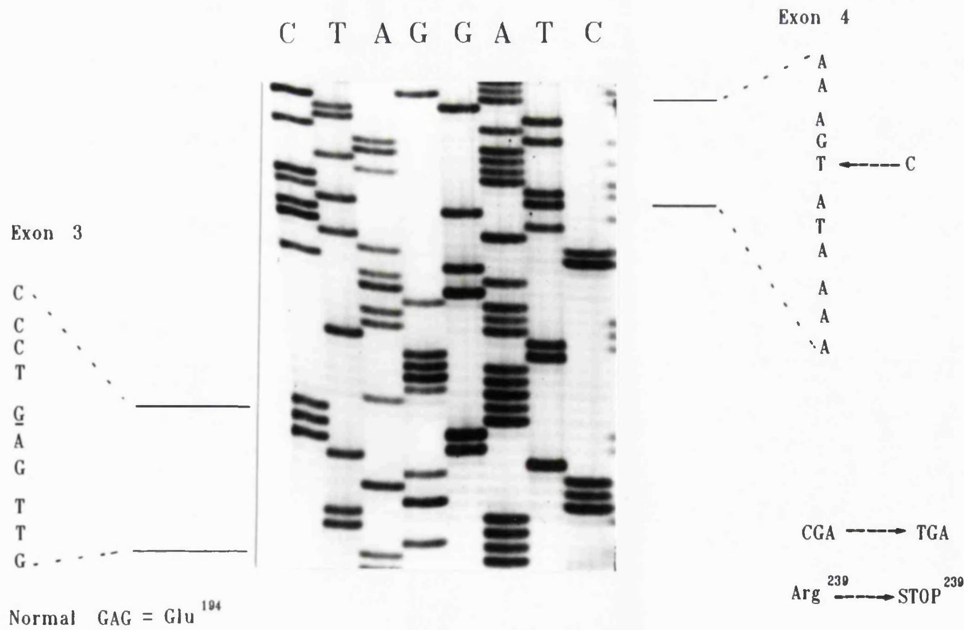


Figure 3.2e: Nucleotide sequences of exon 3 and exon 4 of *CYP17*, obtained by sequencing cloned DNA.



Sequencing of cloned exon 3 and 4 DNA showed a normal exon 3 on the same allele as a mutant exon 4, which carried an R239X mutation.

3.3 Discussion of methodology

Each of the five clones isolated from the gene library were found to contain the 5.7 kb *Eco*RI fragment of *CYP17*. The absence of the 6.9 kb fragment may have been due to preferential uptake and maintenance of smaller fragments of DNA within the lambda gt10 bacteriophage, which has an upper limit of approximately 6 kb. This vector, however, has been used successfully, by other groups, to isolate *CYP17* (Dr. M.R. Waterman, personal communication). Other lambda vectors have the capacity for cloning larger inserts (Kaiser & Murray, 1985). Gene library screening of lambda was employed by Picardo-Leonard & Miller (1987) and Kagimoto *et al* (1988) to isolate the genomic sequence of *CYP17*, although the latter used the EMBL3 lambda vector, which is designed to accept fragments up to 20 kb in size (Frischauf *et al*, 1993).

Several apparent false positives, which hybridized in the initial primary screen, failed to produce positive plaques on secondary screening. This result may have been due to incorrect alignment of blots and plaques when picking clones or insufficient stringency of washing after hybridization. Perseverance with the library screening may have produced further clones, perhaps including exons 4-8.

The process of library screening was time consuming, taking almost a year to isolate the same three exons of *CYP17*. The main advantage of the technique is that large fragments of DNA can be obtained and the genomic sequence isolated in cases where only a cDNA was previously available. The direct cloning of genomic DNA employed by the technique is also of use in cases where PCR amplification proves difficult, as often occurs in the 5' regions of genes, where repetitive sequences can lead to anomalies in PCR amplified DNA. However, as a method of screening genes of known sequence for mutations it was shown to be too labour intensive.

3.4 Discussion of results

The patient had all the classic symptoms of 17α-hydroxylase deficiency. A normal gene arrangement was found on Southern blot analysis excluding any major gene deletion or rearrangement and the cause of the 17α-hydroxylase deficiency was found to be due to the two point mutations located. These mutations would produce a truncated and non-functional 17α-hydroxylase enzyme, due to the lack of haem or substrate binding ability.

Mutations have been described in 23 other cases of 17α-hydroxylase deficiency and range from small duplications and deletions to amino acid substitutions and premature termination codons, as described in section 1.4.3. The mutation reported in this thesis at codon 239 involves a CpG dinucleotide, which are recognised as mutational hotspots (Bird, 1980). Interestingly, one other patient has been described with the same nonsense mutation in exon 4 described here (Ahlgren *et al*, 1992). In that case, however, the other allele had an amino acid substitution of threonine for proline at position 342 in exon 6, caused by the change CCA→ACA. The patient discussed here, who is English, may share a common ancestry with the Canadian patient studied by Ahlgren *et al*, although the two are not believed to be directly related.

4- Screening for mutations in the *CYP11B* genes

4.1 Subjects with mutations in the *CYP11B* genes

DNA samples of twelve individuals with suspected mutant *CYP11B* genes were available for study. The twelve individuals came from 9 unrelated kindreds (Table 4.1). Six kindreds (1-5 and 7) had children affected by 11ß-hydroxylase deficiency, which was presumed to be caused by defects in the *CYP11B1* gene. One kindred (8), of Turkish origin, had two children with combined 11ß-hydroxylase and aldosterone synthase deficiency, presumed to be caused by mutations in both *CYP11B1* and *CYP11B2* genes. Subject 9 (kindred 6), of Pakistani origin, was diagnosed as a case of CMO II deficiency. The final kindred (9), of Greek origin, was comprised of a mother and daughter, both of whom were born with severe hypertension, which was alleviated by adrenalectomy and replacement steroids in the case of the mother. Mutation within the *CYP11B1* or *CYP11B2* gene was a possible cause of disease in this family. In several cases DNA samples from parents and siblings were also available, which increased the number of individuals to be screened to 16.

The above diagnoses were made on the basis of abnormal urinary steroid profiles, determined in the laboratory of Dr. John Honour (Middlesex Hospital, London), and blood hormone measurements obtained in the laboratory of Dr. R. Fraser (Medical Research Council Blood Pressure Unit, Glasgow), following clinical referral for endocrine assessment. In cases of 11ß-hydroxylase deficiency excess excreted THS in the range of 5,000-30,000 μg/ 24h was detected (normal range < 50 μg/ 24 h) and THDOC in the range of 700-5000 μg/ 24 h (normal value of <50 μg/ 24 h) (Schneider, 1991). Urinary androgens (androsterone and aetiocholanolone) were also elevated. Females typically presented at birth with virilized genitalia, while males were detected due to precocious sexual and physical development driven by excess

androgens. Several patients within this group were initially diagnosed as having

Table 4.1: Subjects with suspected mutations in *CYP11B* genes

Kindred	Subject	Sex	Ethnic Origin	Deficiency/ Disease
1	1	Female	Indian (Asian)	11ß-OH
2	2 (Mother)	Female	Indian (Asian)	-
2	3 (Daughter)	Female	Indian (Asian)	11ß-OH
3	4 (Daughter)	Female	Nigerian	11ß-OH
3	5 (Father)	Male	Nigerian	-
4	6	Male	English	11ß-OH
5	7 (Brother)	Male	English	11ß-OH
5	8 (Brother)	Male	English	11ß-OH
6	9	Female	Pakistani	CMO II
7	10	Male	Indian (Asian)	11ß-OH
8	11 (Father)	Male	Turkish	-
8	12 (Mother)	Female	Turkish	-
8	13 (Son)	Male	Turkish	11ß-OH + Aldosterone Synthase
8	14 (Son)	Male	Turkish	11ß-OH + Aldosterone Synthase
9	15 (Mother)	Female	Greek	Severe Hypertension
9	16 (Daughter)	Female	Greek	Severe Hypertension

21-hydroxylase deficiency, which was changed to 11ß-hydroxylase deficiency following the development of hypertension and reevaluation. Subject 4, an affected female, presented at birth with virilization, and was treated for 21-hydroxylase deficiency, but subsequently developed hypertension (140/ 105 mm Hg) following a five year lapse in hydrocortisone treatment between the age of 5 and 11 on returning to Nigeria (personal communication from Dr. P.J.A. Moult, Consultant Endocrinologist, Whittington Hospital, London). The diagnosis of subject 6 was also altered to 11ß-hydroxylase deficiency following the development of hypertension in early childhood.

The clinical details of subject 13 are discussed in Spoudeas *et al* (1993). At age 11 this patient presented with hypertension (220/ 140 mm Hg) causing retinopathy and cardiomyopathy. Elevated androgens, THDOC and THS were discovered on urinary steroid analysis. Dexamethasone treatment reduced the levels of the above steroids, by suppression of ACTH driven steroidogenesis. The treatment relieved the suppression of plasma renin activity (PRA), by reducing the level of excess DOC, but aldosterone synthesis did not resume, even after two years on treatment. This suggested that a defect was also present in aldosterone synthase in this patient, a finding confirmed by Spoudeas *et al* (1993). Subject 14, the younger brother of subject 13, was similarly affected.

The mother and daughter of Greek origin, subjects 15 and 16, were affected by an adrenal related form of hypertension, that was believed to involve a glucocorticoid remediable form of hyperaldosteronism. Recent studies have shown that this condition can be caused by unequal crossover between the *CYP11B* genes (Pascoe *et al*, 1992a; Lifton *et al*, 1992). The mother's symptoms were alleviated by adrenalectomy, confirming that the cause of hypertension was of adrenal origin. Excess 18-hydroxysteroid products (partially suppressible by dexamethasone treatment) were detected in subject 16, which may be the cause of the hypertension.

Subject 9, a patient with reported *hypotension* suspected to be caused by CMO II deficiency, was included in the study, as the processes leading to mutations in the *CYP11B2* gene were also of interest. The possibility that gene conversion and non-homologous recombination between the *CYP11B1* and *CYP11B2* genes leads to loss

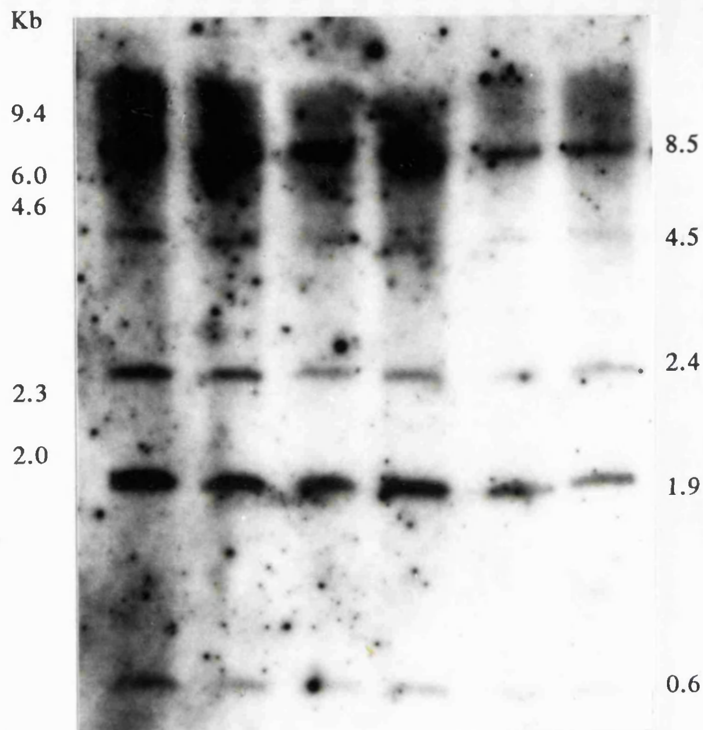
or alteration of enzyme function was examined.

4.2 Strategy for locating mutations

In order to screen the *CYP11B1* and *CYP11B2* genes from the 16 individuals for mutations, a fast and efficient means of locating DNA sequence changes had to be employed. A strategy was devised that began with Southern blot analysis to eliminate the possibility of large scale deletions or insertions being present. The human *CYP11B* genes are highly homologous, but give different band sizes on digestion with *Bam*HI endonuclease (Mornet *et al*, 1989). Southern blot analysis was carried out on the above group of individuals, as described in section 2.2.8, following restriction digestion with *Bam*HI. A radioactive probe was prepared from a *CYP11B1* cDNA fragment, pH11F2, kindly supplied by Dr. P.C. White, New York, and used for hybridization to reveal fragments of both *CYP11B* genes. Five bands, at 8.5 and 1.9 kb (*CYP11B1*) and 4.5, 2.4 and 0.6 kb (*CYP11B2*), corresponding to the normal digestion products of the two *CYP11B* genes, were obtained from the DNA of all subjects studied (Figure 4.2a). This suggested that no large insertion or deletion had occurred in the *CYP11B* genes of the patients listed in Table 4.1.

The symptoms of subjects 15 and 16 were consistent with glucocorticoid remediable hyperaldosteronism. Rehybridization of the Southern blot with a probe prepared from PCR amplified *CYP11B1* exon 4 DNA did not detect an additional 6.3 kb *Bam*HI fragment, formed by fusion of the 5' region of *CYP11B1* with the 3' region of *CYP11B2*, which has been characterized in GRH (Lifton *et al*, 1992). The possibility that a gene conversion event had altered a small region of the *CYP11B1* gene of subjects 15 and 16 to encode amino acids present in CYP11B2 remained a possibility.

Figure 4.2a: Southern blot analysis following *Bam*HI digestion of genomic DNA



Above is a representative Southern blot of DNA from subjects 1, 6, 13 and 14 (lanes 1-4) and two normals (lanes 5-6). In each lane 10 μg of genomic DNA was loaded, alongside lambda *Hind*III size markers. All subjects gave the normal band sizes of 8.5 and 1.9 kb, corresponding to *Bam*HI digestion of *CYP11B1*, and 4.5, 2.4 and 0.6 kb, corresponding to *Bam*HI digestion of *CYP11B2*.

The lack of Southern blot abnormalities suggested that any mutations within the *CYP11B* genes of the subjects were due to single base changes or small deletions and insertions. To locate such small scale changes the strategy shown in Figure 4.2b was employed.

The *CYP11B1* and *CYP11B2* genes were specifically amplified by PCR. Primers were selected which were complementary to sequences of DNA that differed between the two *CYP11B* genes (Appendix 1). These specific DNA fragments were used as template DNA for a second round of PCR amplification, in which individual exons were amplified using primers complementary to both *CYP11B* genes.

Screening for single base changes

Various techniques have been developed to detect single base sequence changes in PCR amplified DNA, these include: RNAse cleavage (Myers *et al*, 1985), chemical cleavage (Cotton *et al*, 1988), denaturing gradient gel electrophoresis (DGGE) (Fischer & Lerman, 1980; Myers *et al* 1987; Sheffield *et al*, 1989; 1992) and heteroduplex analysis (Keen *et al*, 1991, White *et al*, 1992). One of the most widely used techniques, due to its simplicity, is single-strand conformation polymorphism analysis (SSCP) (Orita *et al*, 1989a; 1989b). This technique is based on the principle that single-stranded DNA molecules take on specific, sequence-based, secondary structures (conformers) under nondenaturing conditions. Molecules differing by even a single base substitution may form different conformers and migrate differently in a non-denaturing polyacrylamide gel (Figure 4.2c). This technique has proven useful for detection of multiple mutations and polymorphisms (Demers *et al*, 1990; Mashiyama *et al*, 1990; Orita *et al*, 1990; Suzuki *et al*, 1990; Murakami *et al*, 1991; Reiss *et al*, 1992).

SSCP has been reported to be most sensitive when used to detect sequence variation in PCR amplified DNA molecules 200 bp or less in size (Hayashi, 1992). Some evidence suggests that detection of single base substitutions is less sensitive in larger

Figure 4.2b: Strategy for screening *CYP11B* genes for mutations

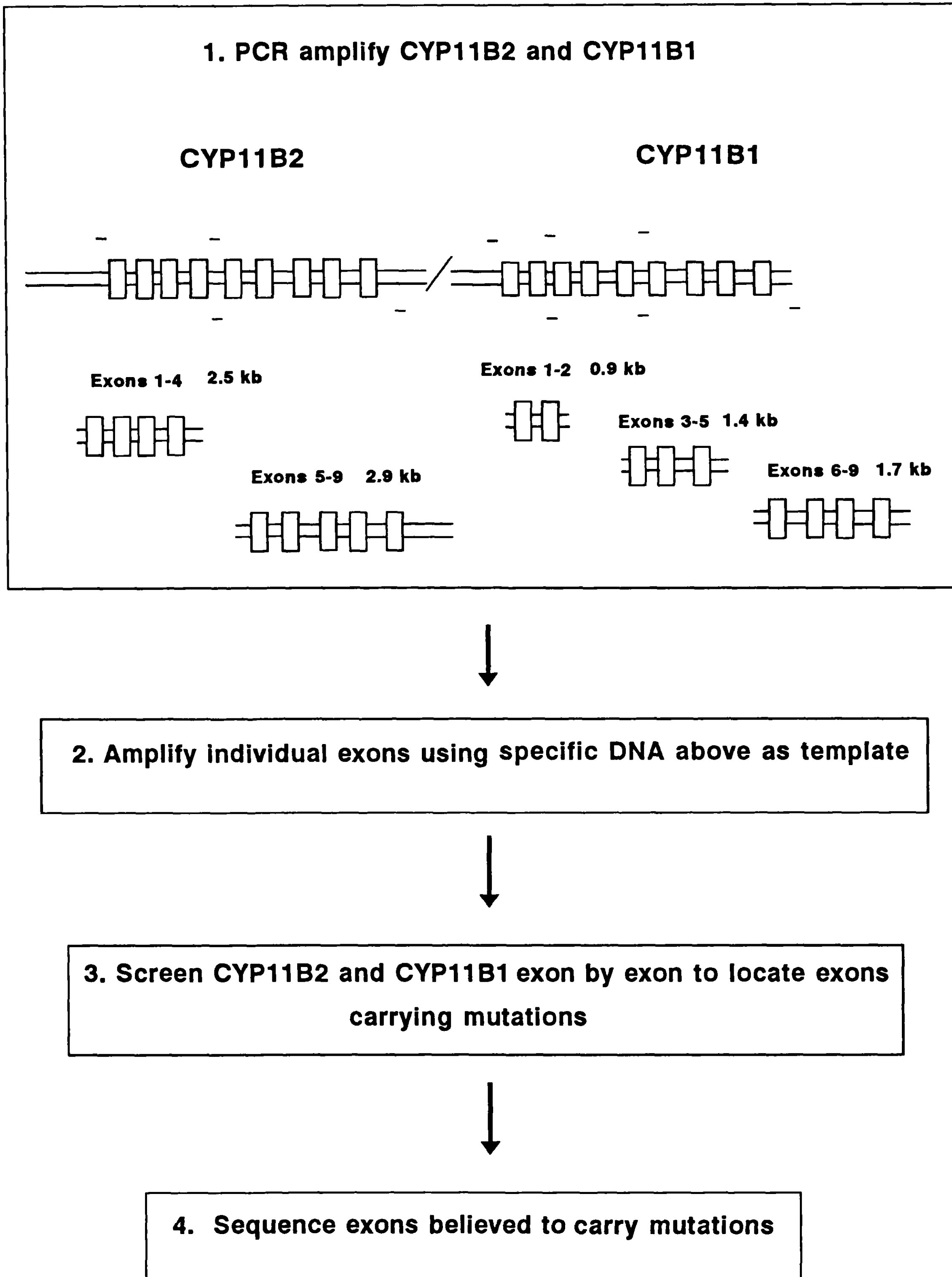


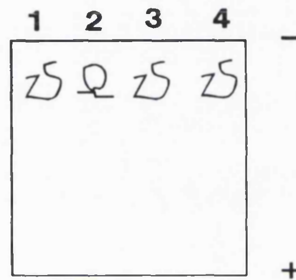
Figure 4.2c: Single strand conformation polymorphism (SSCP) technique for detecting single base changes within PCR amplified DNA fragments.

1 PCR 250bp DNA fragments using radioactive ^{32}P dCTP.

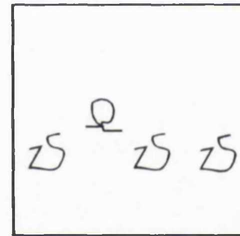
A single exon can be amplified from several individuals.

2 Single-strand DNA by addition of formamide and heating at 95°C.

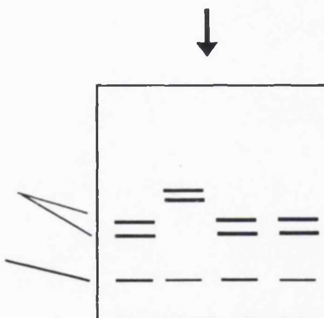
3 Electrophorese fragments on a 6% non-denaturing polyacrylamide gel.



Differences in base sequence may affect conformation of single-strand DNA as it moves through the gel and this may cause a band shift.



4. Autoradiograph for 12-72 hours.



SSCP analysis relies on the ability of single stranded DNA to adopt a sequence-specific conformation as it migrates through a non-denaturing polyacrylamide gel: altered sequence can give rise to altered conformation and altered migration rate. This can reveal which individuals carry sequence changes relative to normal controls. The DNA bands are visualized by autoradiography and sequence changes give band shifts as shown above.

DNA fragments (Sarkar *et al*, 1992). In a systematic study of the limitations and strengths of SSCP carried out by Sheffield *et al* (1993), the optimal size fragment for sensitive base substitution detection by SSCP was reported to be 150 bp, with a reported 97% detection rate of mutations within the mouse globin promoter. It was also reported that some base substitutions in particular pieces of DNA may not affect conformation, that different mutations at the same position may give indistinguishable band shifts and that G to T transversions are detected at a lower rate than all other mutations (57% as opposed to 79%). Detection of mutations can be increased by the use of different primers to amplify the region of interest, or by digestion of larger fragments to generate DNA in the range of 150 bp. Thus SSCP analysis is a useful method of primary screening for mutations and common polymorphisms, but does not have 100% sensitivity. The SSCP technique was used to screen *CYP11B1* and *CYP11B2* genes of the subjects listed, exon by exon, to locate regions carrying sequence changes.

4.3 SSCP analysis

4.3.1 Radioactive SSCP technique

The technique used initially was a modification of that developed by Orita (1989a). DNA specific to either *CYP11B1* or *CYP11B2* was prepared in 25 μl reactions containing 1 μM each primer (Appendix 1), 1.5 mM $MgCl_2$, 0.2 mM dNTPs and 1 U Taq polymerase in 1x reaction buffer (Promega, Southampton, UK). *CYP11B1* was amplified in three fragments of 0.9, 1.4 and 1.7 kb, and *CYP11B2* in two fragments of 2.5 and 2.9 kb, as shown in Figure 4.2b. The thermal cycle programme was as for standard PCR, see section 2.2.4, with the exception that the extension time was increased to 2.5 minutes at 72°C and the annealing temperature lowered to 56°C.

The amount and purity of DNA produced was often poor, particularly in the case of the larger products. To overcome this problem the products were electrophoresed on 0.8% agarose gel in 1x TBE. The required fragment was excised from the gel, added to 200 μl of sterile ddH_2O and heated to 98°C for 15 minutes to melt the agarose and make single stranded DNA template at a concentration of approximately 0.05 ng/ μl, which was used in the second round of amplification described overleaf.

SSCP-PCR reactions, for each exon, were prepared in 10 μl total volumes containing: 0.15 ng purified specific DNA template, 1 μM each primer, 1.5 mM $MgCl_2$, 0.2 mM dATP, dTTP, dGTP, dCTP, 1 μCi ^{32}P-dCTP (3000 Ci/ mmol, NEN, Stevenage, UK), 0.25 U Taq polymerase in 1x reaction buffer. Amplification conditions were as for standard PCR (see section 2.2.4). The final PCR product was added to 10 μl of a 95% formamide, 20 mM EDTA, 0.05% bromophenol blue solution. Samples were denatured by heating at 95°C for 3 minutes, then 4 μl of each loaded onto a 20x30 cm, 0.4 mm thickness, 6% acrylamide gel (19:1 acrylamide:bisacrylamide), 10% (v/ v) glycerol and 1x TBE for sequencing buffer. The samples were electrophoresed at 30 W (1875 V and 16 mA) at room temperature. Alternative SSCP electrophoresis conditions used for comparison were as above except at 4°C, or as above at room temperature with no glycerol in the gel. Following electrophoresis the gel was dried and exposed to Fuji X-100 X-ray film for 12-72 hours, to visualize the radioactive DNA bands.

4.3.2 Factors that improved SSCP resolution

A number of improvements to the initial SSCP technique were made, following suggestions made by Hogg *et al* (1992). Initially the amount of cold dCTP in the PCR reaction was at 0.2 mM concentration. By reducing this ten fold to 0.02 mM the incorporation of radioactive ^{32}P-dCTP was increased. In the unmodified method above, 20-40 ng of each DNA sample was loaded on the gel. After reducing the concentration of dCTP to 0.02 mM, only 2-5 ng of radioactive PCR product was required for visualization of bands (Figure 4.3). Instead of a 1:1 dilution of PCR DNA with loading buffer, a 1:25 dilution was made: 2 μl of PCR product was diluted in 50 μl of formamide/ bromophenol blue loading buffer and 6μl loaded per well. The reduction in DNA improved the detection of sequence specific differences in band pattern.

Addition of SDS at 0.03% and EDTA at 16 mM final concentration to the DNA samples before denaturing improved resolution of SSCP bands, presumably by reducing both inter and intrastrand associations that could give alternative

conformations. Bands became much sharper and there were less problems with multiple, non-specific band patterns.

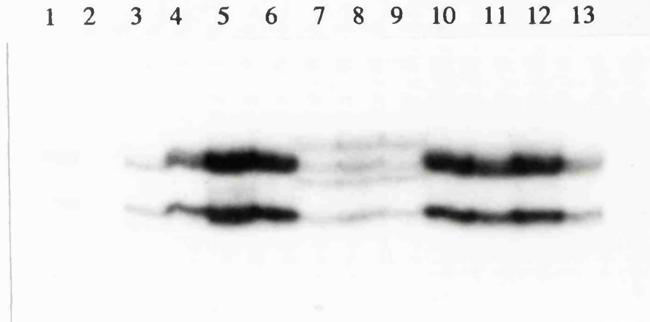
A minor improvement was made by changing the comb used to prepare the wells of the gel. When sharks' tooth combs were used the bands were 3 mm in length, bowed at either end and continuous with adjacent samples (Figure 4.3). A crenellated comb, that produced 6 mm wells, spaced 4 mm apart, allowed a better comparison of bands with less distortion.

4.3.3 Preparation of DNA samples for SSCP analysis

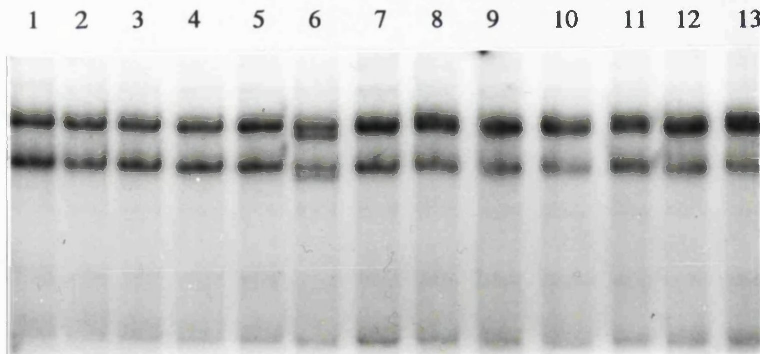
In the improved method, a 2μl aliquot of the final SSCP-PCR product was added to 16 μl of 0.1% SDS, 10 mM EDTA, and 32 μl of 95% formamide, 20 mM EDTA, 0.05% bromophenol blue loading buffer. Samples were denatured by heating at 95°C for 3 minutes, the samples immediately placed on ice, then 6 μl of each sample electrophoresed as detailed in Table 4.2.

If the PCR product was greater than 250 bp in length, digestion with restriction endonuclease enzymes was carried out to give fragments of the desired size. A 2 μl aliquot of PCR product was added to 3 μl of a master mix, which contained 0.5 μl of endonuclease (5-6 U total), 0.5 μl of 10x endonuclease buffer, 2 μl ddH_2O, and digestion carried out overnight at 37°C or optimal temperature for enzyme digestion. The digested samples were then prepared as described above. Electrophoresis running times were adjusted to prevent small DNA fragments from running off gels. Table 4.2 gives the size of digested DNA fragments for the different exons of *CYP11B1* and *CYP11B2* and the approximate running times for the different gels used.

Figure 4.3: Improvement of original SSCP technique.



In the original method (section 4.3.1) SSCP bands were diffuse, and differences unclear. The samples above are of a 306 bp PCR fragment of exon 4 of *CYP11B1* (see Table 4.2), with 20-40 ng of ^{32}P-dCTP labelled DNA in each lane. Lanes 7-9 contain DNA of *CYP11B2* and *CYP11B1*, prepared by PCR from genomic DNA, whilst the other lanes contain specifically amplified *CYP11B1* DNA. Lane 5 contains DNA of subject 6 and shows a faint extra band.



The same region of DNA was subjected to SSCP analysis with the improved technique (section 4.3.2). Visualization of bands with only 2-5 ng of DNA was possible, due to higher incorporation of ^{32}P-dCTP. The bands were less diffuse than in the original method and subject 6 (lane 6) showed clear doublet bands, indicating a sequence difference. The gels and electrophoresis conditions used for the original SSCP analysis and the improved SSCP analysis were identical (6% polyacrylamide, 19:1 acryalmide: bisacrylamide, in 1x TBE, with 10% glycerol, run at for 3 hours at 60W at 4°C).

Table 4.2: PCR amplification fragment sizes of *CYP11B1* and *CYP11B2*, endonuclease digestion required for SSCP analysis and running times on SSCP gels

Exon	Primer pairs*	Size (bp)	Digestion	Fragments (bp)	Gel running times (hours) GRT	G4°C	N-GRT
CYP11B1 1	072T-7W	388	Pst I	165, 223	4.5	3.5	2.0
2	8W-073T	260	Rsa I	77, 183	4.0	3.0	1.5
3	9W-10W	349	Hha I	115, 234	4.0	3.0	1.5
4	11W-12W	306	Stu I	210, 96	4.0	3.0	1.5
5	H1042-075T	312	Rsa I	123, 189	4.0	3.0	1.5
4+5	11W-075T	734	Rsa I	330, 215, 189	5.0	4.0	2.0
6	076T-13W	224	-	-	5.0	4.0	2.0
7	14W-15W	147	-	-	4.0	3.0	1.5
8	371Y-A617	262	Hae III	119, 143	4.0	3.0	1.5
9	A618-077T	224	-	-	5.0	4.0	2.0
CYP11B2 1	709V-7W	487	Hha I	199, 288	5.0	4.0	2.0
2	8W-073T	260	Rsa I	77, 183	4.0	3.0	1.5
3	9W-10W	349	Hha I	116, 233	4.0	3.0	1.5
4	11W-12W	306	Stu I	210, 96	4.0	3.0	1.5
5	711V-369Y	337	Rsa I	148, 189	4.0	3.0	1.5
6	370Y-13W	224	-	-	5.0	3.0	2.0
7	14W-15W	147	-	-	4.0	3.0	1.5
8	371Y-A617	262	Hae III	119, 143	4.0	3.0	1.5
9	A618-16W	403	Stu I	78, 100, 210, 15	4.0	3.0	1.5

In the table above:

GRT= 6% polyacrylamide gel (19:1 acrylamide:bisacrylamide) in 1x TBE, containing 10% glycerol electrophoresed at 30 W (1875 V, 16 mA) at room temperature.

G4°C= 6% polyacrylamide gel (19:1 acrylamide:bisacrylamide) in 1x TBE, containing 10% glycerol electrophoresed at 60 W (1875 V, 32 mA) at 4°C (in a cold-room).

N-GRT= 6% polyacrylamide gel (19:1 acrylamide:bisacrylamide) in 1x TBE, containing no glycerol electrophoresed at 30 W (1875 V, 16 mA) at room temperature.

All gels were pre-warmed by running at the prescribed power for 15 minutes before loading samples.

* Pairs of primers used in PCR amplification were assigned arbitrary names (defined in Appendix 1).

4.4 Variations of SSCP conditions

When SSCP was first evaluated 20x30 cm gels with 0.4 mm spacers were used on a standard sequencing rig. When the technique was altered larger gels, 40x30 cm with 0.5 mm spacers were used to accommodate the increased well size and spacing of samples.

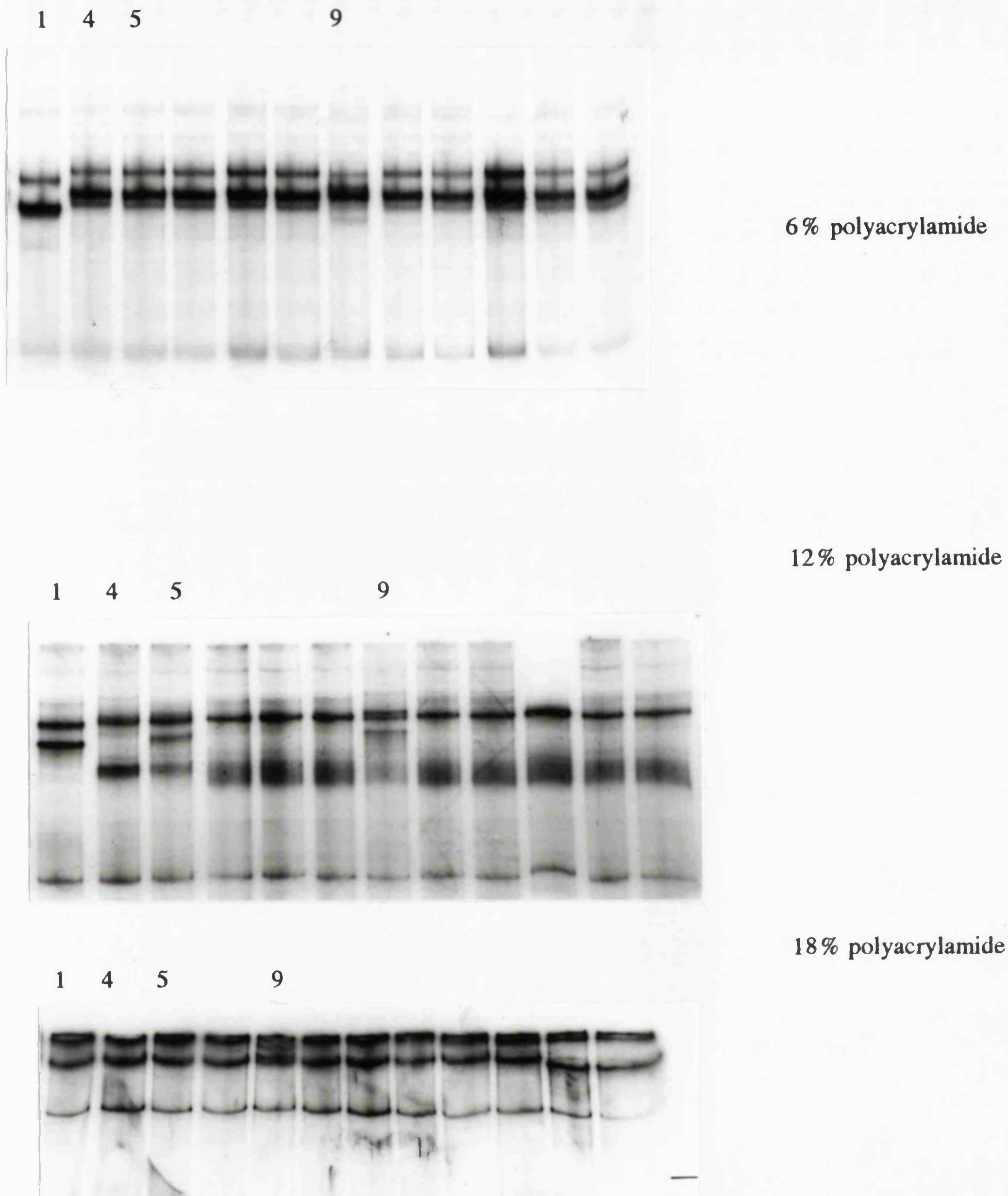
A comparison of different concentrations of polyacrylamide was also carried out. SSCP analysis of *CYP11B1* exon 6, using a 6% polyacrylamide gel, containing 10% glycerol identified only two samples to be markedly different. When the polyacrylamide gel concentration was increased to 12% (still with 10% glycerol) two more samples gave unique SSCP patterns. However, the bands appeared compressed relative to those obtained on the 6% gel. At 18% polyacrylamide and 10% glycerol the unique patterns were no longer resolved, the radioactive DNA bands being very thin and compressed and migration distance only several cm after 5 hours electrophoresis (Figure 4.4a).

It appeared that 12% polyacrylamide (19:1 acrylamide: bisacrylamide) with 10% glycerol was nearing the useful upper limit of polyacrylamide concentration for obtaining different DNA conformations by SSCP analysis. For general screening 6% polyacrylamide (19:1 acrylamide: bisacrylamide) with or without 10% glycerol, was used.

Temperature and gel composition greatly affected SSCP patterns. The inclusion of glycerol in gels retarded the migration of DNA and increased conformational differences. The least satisfactory results were obtained with non-glycerol gels, electrophoresed at 4°C and 60 W. These electrophoresis conditions gave the fastest rate of DNA migration which may have reduced differences in conformation. Non-glycerol gels at room temperature and 30 W, glycerol gels at room temperature and 30 W and glycerol gels at 4°C and 60 W all proved useful in the SSCP analysis. Comparison of the three different conditions allows the following conclusions to be drawn: non-glycerol gels at room temperature produced sharp bands, but less marked differences were seen; glycerol gels at room temperature produced rather diffuse

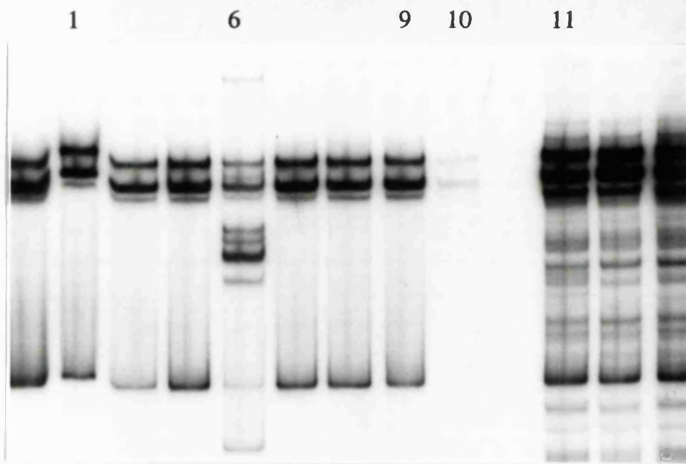
bands, but differences could be seen; glycerol gels at 4°C had the advantage of generating sharper bands than at room temperature, while maintaining the differences in migration pattern. In practice it was useful to carry out all three types of conditions for each exon studied, as differences just visible on a non-glycerol gel were often clearer on glycerol gels at the two temperatures, whilst in the cases where the bands on glycerol gels were diffuse the non-glycerol gel often showed a clearer difference in pattern. A comparison of results from the different SSCP conditions is shown in Figure 4.4b.

Figure 4.4a: Effect of varying polyacrylamide concentrations on SSCP patterns

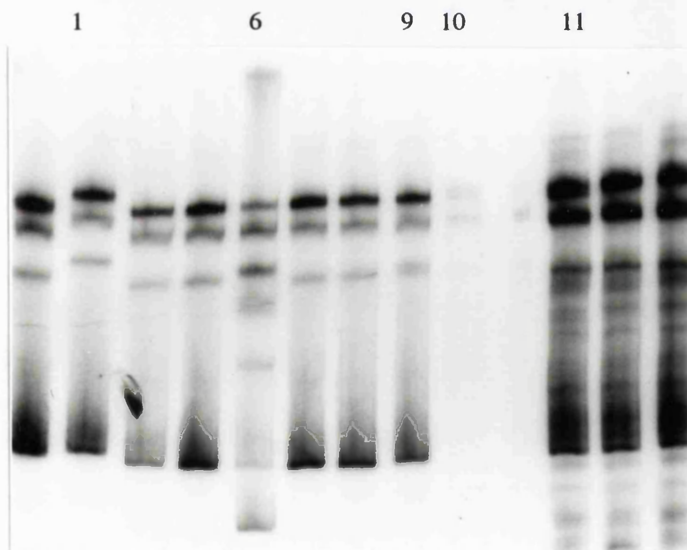


SSCP analysis of a 224 bp PCR fragment, including *CYP11B1* exon 6, was carried out on gels of 3 different concentrations of polyacrylamide: 6%, 12% and 18% (19:1 acrylamide: bisacrylamide). All gels contained 10% glycerol and 1x TBE, and were run for 5 hours at room temperature. Subjects 1 and 9 show SSCP differences on the 6% gel. Subject 1, 4, 5 and 9 show differences on the 12% gel, while only subject 9 shows a difference on the 18% gel. Numbers above lanes represent individuals described in Table 4.1, p.121.

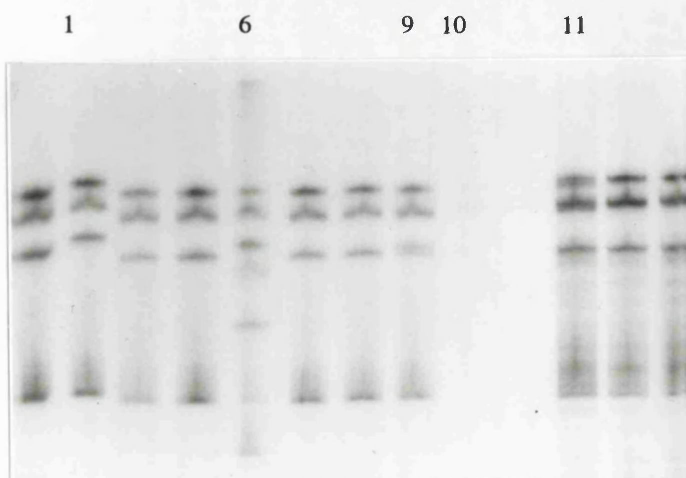
Figure 4.4b: Effect of glycerol and temperature on SSCP patterns of *CYP11B1* exon 2.



Non-glycerol gel, RT. Bands sharp, size differences shown in subjects 1 and 6. Subject numbers are given above corresponding lanes for samples, as per Table 4.1, p.121.



Glycerol gel, RT. Bands more diffuse, but differences in DNA of subjects 10, 9 and 11.



Glycerol gel, 4°C Bands sharper than for glycerol, RT. SSCP differences for 9 and 11 confirmed.

4.5 Development of non-radioactive SSCP analysis using silver staining

4.5.1 Non-radioactive SSCP analysis

Radioactive SSCP analysis produces a large amount of low grade radioactive waste and involves exposure of the operator to radiation for lengthy periods of time, due to the manipulation of many samples through a number of processes. In an effort to reduce the use of radioactive material and exposure to radiation, the use of silver staining to visualize DNA was investigated. The technique of Merrill *et al* (1981) was used, as described in section 2.2.6. The SSCP PCR conditions were the same as described for radioactive samples, except for the omission of ^{32}P-dCTP and increasing dCTP to 0.2 mM. The sensitivity of silver staining was found to be far less than radioactive labelling when using a 40x30 cm gel. A 6 μl aliquot of PCR product (around 30-60 ng) was added to a solution containing 6 μl 0.8% SDS/ 80mM EDTA, 6 μl of a formamide/ bromophenol blue loading buffer and 2 μl 50% glycerol. All 20 μl of this mixture was denatured by heating at 95°C for 3 minutes, immediately quenched on ice, and electrophoresed as above. The three different conditions of electrophoresis described above were carried out using non-radioactive PCR DNA.

4.5.2 Improving silver staining SSCP technique

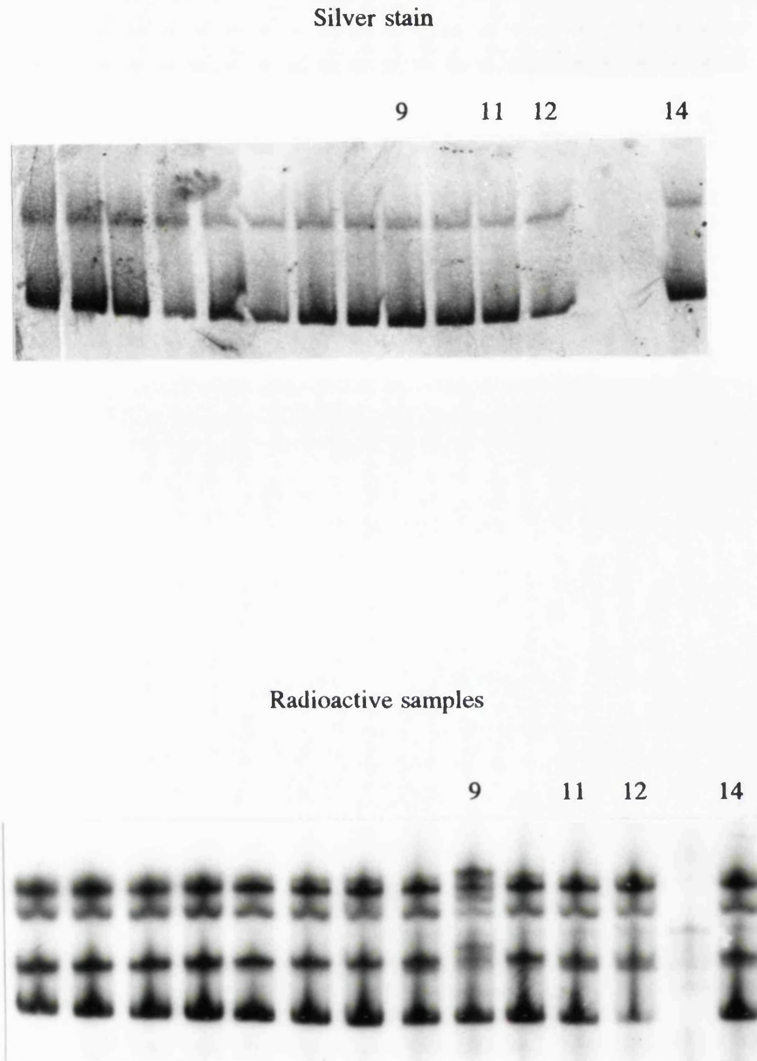
Multiple bands were sometimes observed on silver stained gels, as a result of increased amount of DNA loaded. Excess DNA leads to association of single strands or incomplete denaturation on heating, so that multiple conformers are produced that affect SSCP patterns. Attempts were made to improve the silver staining so that visualization of around 5 ng of DNA could be achieved.

One problem with silver staining was high background staining. In part this may result from carry over of staining solutions between the different steps. The carry over effect was reduced by decreasing the size of the gel. Gels of 20x20 cm x 0.5 mm and 8x8 cm x 0.5 mm size were prepared, with the same composition as given above (6% polyacrylamide, 10% glycerol gels). These were used for SSCP

analysis of exon 2 of *CYP11B1* and silver staining carried out as above. The 20x20 cm gels were run at 8 W (250 V, 32 mA) for a maximum of 4 hours and 2-3 μl of PCR product (10-30 ng DNA) gave an adequate staining of bands. The 8x8 cm gels were run at 0.8 W (160 V, 5 mA) for around 2 hours. As little as 0.5-1.0 μl of PCR product (2.5-10 ng DNA) could be visualized on the 8x8 cm gels. DNA concentration was determined by comparison of PCR product with ϕX174 *Hae*III markers on 0.8% agarose electrophoresis, as described in section 2.2.4. The silver staining thus proved sufficiently sensitive to detect the required amount of DNA.

The resolution of the smaller gels, however, was not as good as the radioactive 40x30 cm gels. Electrophoresis at a controlled temperature and use of more concentrated gels may have overcome the problem of poor resolution. A comparison of silver stained and radioactive SSCP is shown in Figure 4.5.

Figure 4.5: Comparison of SSCP patterns obtained by silver staining with SSCP patterns visualized by radioactivity (for exon 7 of *CYP11B1*).



For silver staining 30 ng of DNA was loaded on each lane. Most of the DNA remained double stranded, so that the single strand bands were faint and no differences were observed. The radioactive technique required 2-5 ng of DNA for visualization and showed subjects 9, 11, 12 and 14 to have SSCP differences.

* A polymorphism is strictly defined as a DNA sequence variant that is present in at least 1% of a population, this being the accepted minimal level for maintenance of a sequence variant within a population of a significant size (greater than 100 individuals).

4.6 Results of SSCP analysis of *CYP11B1* genes

4.6.1 SSCP differences in *CYP11B1*

The number of different SSCP patterns detected for each exon of *CYP11B1* is shown in Table 4.6.1. Many of the patterns were shared by individuals from 2 or more kindreds out of the nine studied, which suggests that these particular variants represent polymorphisms in the general population*: in the following section these presumed polymorphisms are referred to as common variants. Other patterns were unique to affected individuals of a single kindred and a number of these may reflect sequence changes causing loss of CYP11B1 function: in section 4.6.3 these SSCP patterns are termed unique SSCP patterns.

Table 4.6.1: Numbers of different SSCP patterns obtained for each exon of *CYP11B1*.

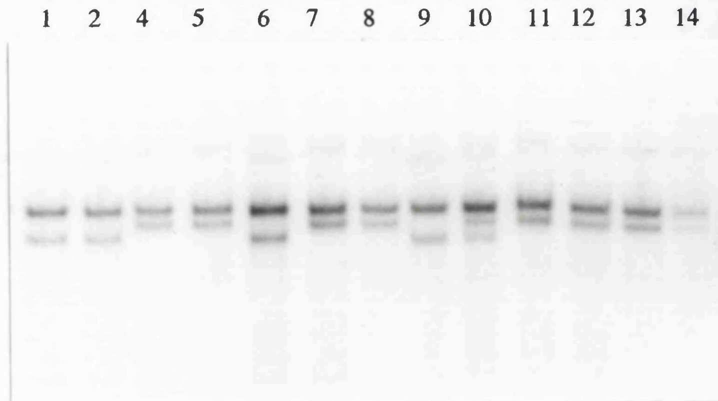
	Exon								
	1	2	3	4	5	6	7	8	9
Different SSCP patterns	3	6	5	5	4	5	3	2	4

4.6.2 The *CYP11B1* gene is polymorphic

DNA from the subjects studied could be divided into two types, based on differences in exons 1, 3 and 4 of *CYP11B1* (Figures 4.6a-c). Samples from individuals 1, 2, 6 and 9 gave the same SSCP patterns for the amplified regions of exons 1, 3 and 4, and were denoted Type A. DNA from subjects 4, 5, 7, 8, 11, 12, 13 and 14, gave a second common variant pattern for exons 1, 3 and 4 and this DNA was denoted Type B. DNA from subjects 3 and 10 produced a combination of both patterns for each of the three polymorphic exons and was defined as Type AB, which was presumed to be a heterozygous combination of A and B alleles. The patterns were not confined to any particular ethnic group - the Type A group including English and Asian subjects, while the second group included Nigerian, Turkish and English subjects.

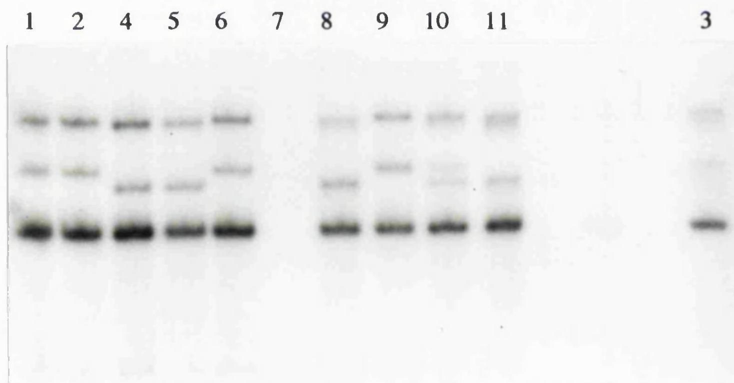
Figures 4.6a-c: Polymorphisms in exons 1, 3 and 4 of *CYP11B1*

a. SSCP of exon 1 of *CYP11B1*, following *Pst*I digest of a 388 bp PCR product.



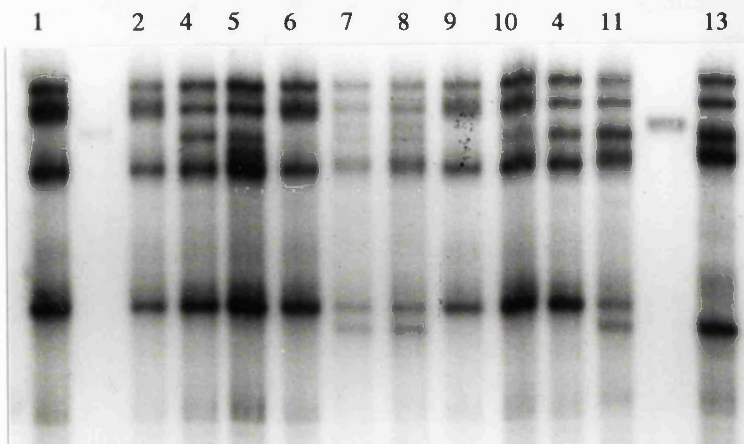
The 165 bp fragment containing 3' of exon 1 is shown. Samples show two polymorphic groups. Subject 10 is heterozygous for the two patterns.

b. SSCP of exon 3 of *CYP11B1*, following *Hha*I digest of a 349 bp PCR product.



The 234 bp fragment, containing the 3' of exon 3 and part of intron 3, is shown. The same two polymorphic groups are shown and subjects 10 and 3 appear heterozygous, although 3 may have additional differences.

c. SSCP of exon 4+5 of *CYP11B1*, following *Rsa*I digest of 734 bp PCR product.



The 330, 215 and 189 bp fragments are shown, containing exons 4-5. There are two common variants of the 330 bp fragment, which divide the subjects into the same two polymorphic groups as for exon 1 and exon 3. This fragment contains intron 4.

Partial sequence analysis, carried out on DNA from individuals of the Type B group, revealed a silent mutation in exon 1 (L75L CTG→CTA) and small gene conversions from *CYP11B1* to *CYP11B2* sequence in the introns just 3' of exons 3 and 4: ATGGGTTG→GTGGGAAG and GGACACCATAGGGCCCCC → GGGACACCATGGGGGCCC respectively, see chapter 5 for details. At these points the sequence differed from that published in the Human Genome Project databank (see Appendix 4 and 5), suggesting that these changes represent the common variants in the regions of exons 1, 3 and 4.

4.6.3 Unique SSCP patterns were found for all affected individuals

For each affected individual unique SSCP patterns were also evident, as shown in Table 4.6.3. The unique patterns were verified by repeat SSCP analysis from genomic DNA. Affected individuals from the same kindred shared the same abnormal SSCP patterns, presumably being of the same genotype. In SSCP analysis of *CYP11B1* exon 5, subject 11 gave bands that appeared to be a mixture of the unique pattern obtained from DNA of his son, subject 13, with a more common variant as would be expected for a heterozygous carrier (Figure 4.6c).

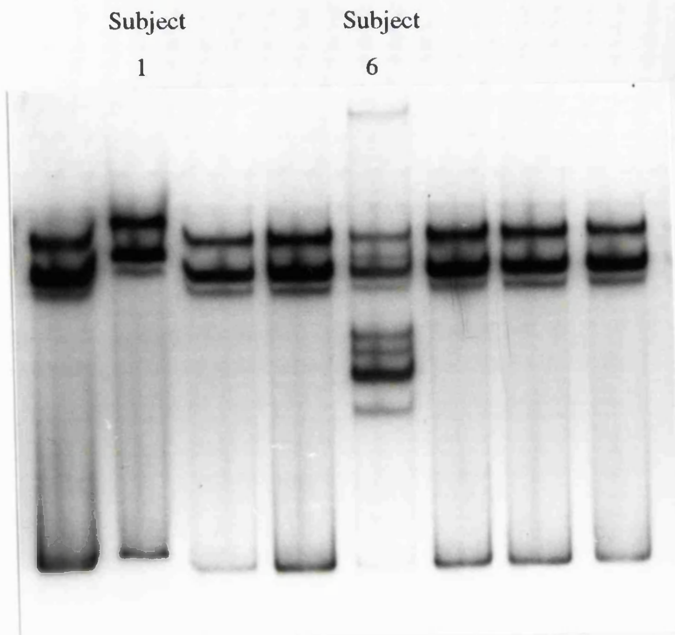
The unique SSCP patterns in exon 2 of DNA from subjects 1 and 6 suggested the presence of a size difference in the amplified DNA (Figure 4.6d). Subject 1 appeared to have an insertion, the DNA migrating at much slower rate than normal, whereas DNA from subject 6 had a deletion on one allele. The size differences could also be seen in the double stranded DNA migrating below the single strand patterns. Subject 1 also had unique patterns in exon 6 (Figure 4.6f), 8 (Figure 4.6g) and 9 (data not shown), while the DNA for exon 4 of subject 6 gave a unique doublet (Figure 4.6e).

All regions that gave unique SSCP patterns were sequenced to ascertain the change in base sequence present, with the exception of exons 2, 6 and 7 from subject 9, a case of CMO II deficiency not believed to carry a deleterious mutation within *CYP11B1*. Sequencing results are given in chapter 5.

Table 4.6.3: Exonic regions of *CYP11B1* giving a unique SSCP pattern for affected individuals

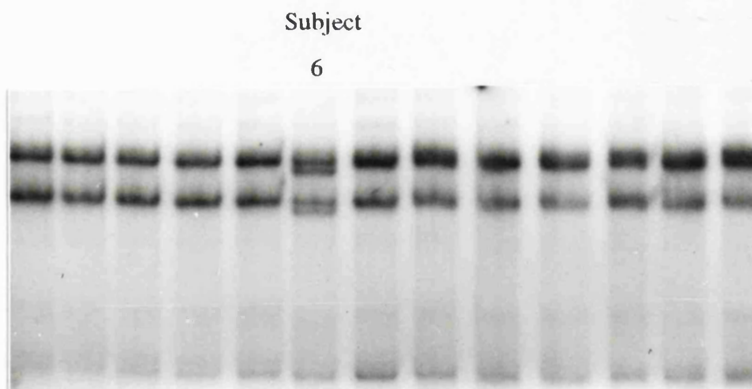
Affected subject	Exonic regions of *CYP11B1* giving unique SSCP pattern
1	2, 6, 8, 9
3	3
4	6
6	2, 4
7 (brother of 8)	5
8 (brother of 7)	5
9	2, 6, 7
10	2, 9
13 (brother of 14)	2, 5, 7, 9
14 (brother of 13)	2, 5, 7, 9
15	4
16	4

Figure 4.6d: SSCP of exon 2 of *CYP11B1*



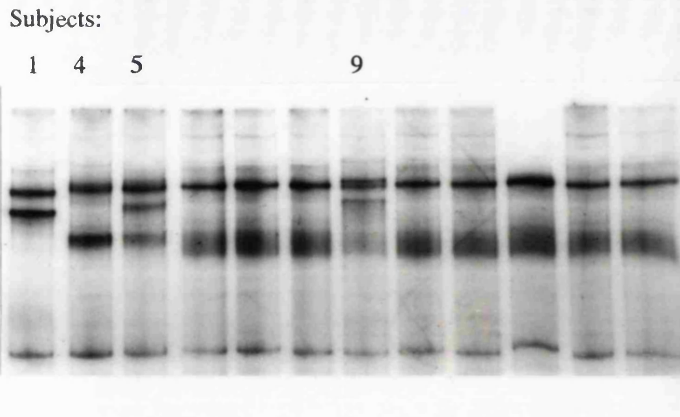
Subject 1 appeared to be homozygous for an insertion, whilst subject 6 appeared to be heterozygous for a deletion. This was shown best by a non-glycerol gel at room temperature (see p. 132 for conditions).

Figure 4.6e: SSCP of exon 4 of *CYP11B1*



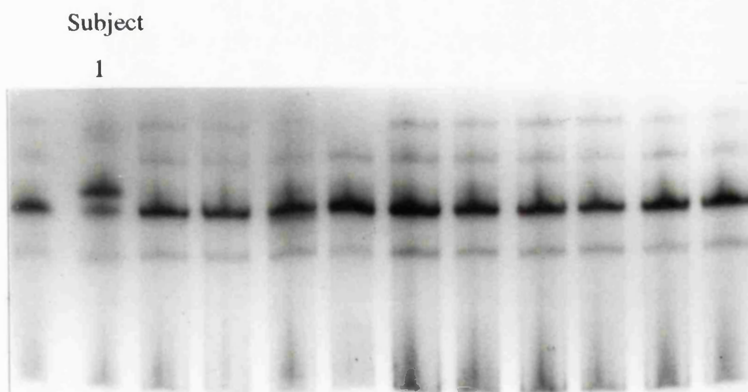
Subject 6 was shown to have a sequence variant generating an SSCP difference for exon 4. The gel used was a 10% glycerol gel at 4°C (see p. 132 for running conditions).

Figure 4.6f: SSCP of exon 6 of *CYP11B1*



Samples from subjects 1, 4, 5 and 9 gave SSCP differences for exon 6. The gel used was a 12% polyacrylamide/ 10% glycerol gel in 1x TBE run at room temperature for 5 hours at 30 W (1875 V, 16 mA).

Figure 4.6g: SSCP of exon 8 of *CYP11B1*



The sample from subject 1 was shown to have an SSCP difference for exon 8. The gel used was a 6 % polyacrylamide/ 10 % glycerol gel in 1x TBE, run at 4°C (for running conditions see p. 132).

4.7 Results of SSCP analysis of *CYP11B2* genes

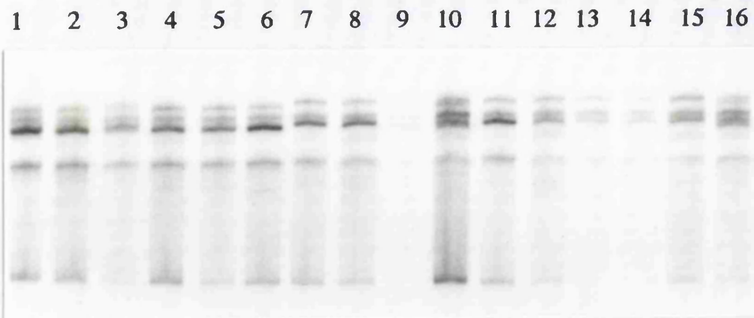
4.7.1 The *CYP11B2* gene is polymorphic

Eight of the nine exons of the *CYP11B2* gene were analyzed by the SSCP technique, exon 1 proving difficult to amplify with the primers available. The *CYP11B2* gene was also found to be polymorphic. Two common variants of *CYP11B2* exon 2 were revealed (Figure 4.7a). Samples from individuals 1, 2, 4, 5, 6 and 9 gave one shared pattern, whilst those of subjects 7, 8, 11, 12, 13, 14 and 15 gave another. This grouping differs from that observed for exons 1, 3 and 4 of *CYP11B1*, in that subjects 4 and 5 shared common patterns with the second group for *CYP11B1*. As before, however, subject 3 and subject 10 were apparent heterozygotes for the two variants of *CYP11B2* exon 2. Subject 16 (not analyzed for variation in exons 1 and 3 of *CYP11B1*) also appeared to to be heterozygous for the two variants of *CYP11B2* exon 2.

Two common variants were also observed for the amplified regions of exons 5 and 6, but there were in addition unique variants, which meant that there were no longer two clearly distinguishable groups. The sequence changes giving rise to the SSCP variants in exons 2, 5 and 6 of *CYP11B2* have not yet been determined.

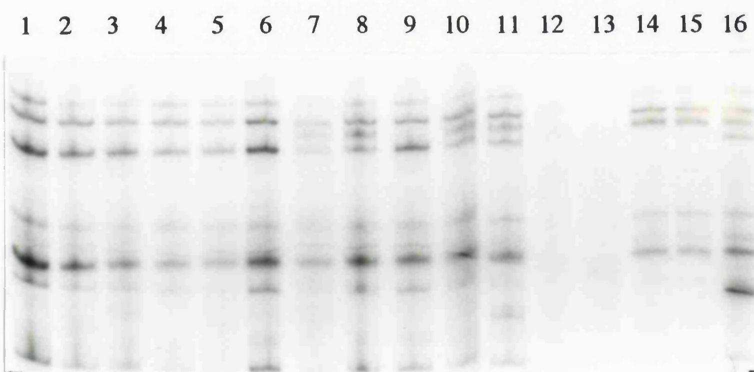
SSCP analysis of exon 9 of *CYP11B2* showed a common variant shared by subjects 12, 13, 14, and 15 (subjects of Greek and Turkish origin), as shown in Figure 4.7b. Sequence analysis of subjects 13, 14 showed the exon 9 variant to be due to a small gene conversion from *CYP11B2* to *CYP11B1* in the intron 5' of exon 9 (see chapter 5 for details). Subjects 7, 8, 11, 10 and 16, appeared to be heterozygous for this variant. Table 4.7.1 gives the summary of SSCP results for *CYP11B2*.

Figure 4.7a: SSCP variants in exon 2 of *CYP11B2*.



The 260 bp fragment containing exon 2 is shown. Two polymorphic groups were revealed. Subjects 1, 2, 4, 5 6 and 9 form one group, and subjects 7, 8, 11, 12, 13, 14 and 15 form another. Subjects 3, 10 and 16 are heterozygous for the two patterns. This grouping differs from that observed for *CYP11B1* exons 1, 3 and 4 (see text for details). The gel was 6% polyacrylamide/ 10% glycerol and was run at 4°C.

Figure 4.7b: SSCP variants in exon 9 of *CYP11B2*



The 210 and 100 bp fragments of a *Stu*I digest of exon 9 are shown. A polymorphism that subdivided Type B individuals (see text for definition of Type B) was revealed. Subjects 14 and 15 are homozygous for the novel pattern, while Subjects 7, 8, 10 and 11 are heterozygous for the novel pattern and the other, more common, variant. Subjects 14 and 15 are of Greek and Turkish origin. Gel conditions were as above.

Table 4.7.1: Numbers of different SSCP patterns obtained for each exon of *CYP11B2*

	Exon								
	1	2	3	4	5	6	7	8	9
Common variants	-	2	1	1	2	2	1	1	2
Heterozygous combination of common variants		1	0	0	0	1	0	0	1
Unique SSCP patterns	-	0	3	1	3	1	1	1	0
Total different SSCP patterns	-	3	4	2	5	4	2	2	3

4.7.2 Unique SSCP patterns from the *CYP11B2* gene

Unique SSCP patterns were seen in exons 3, 4, 5, 6, 7, 8 and 9 of *CYP11B2*, as detailed in Table 4.7.2. These were not confined to individuals with suspected defects in aldosterone biosynthesis and consequently may reflect non-pathological changes. The regions of *CYP11B2* giving unique patterns for the three subjects believed to have aldosterone synthase deficiency were sequenced and the details are given in the following chapter.

Table 4.7.2: Distribution of unique SSCP patterns in *CYP11B2*

Subject	Exonic regions of *CYP11B2* giving unique SSCP pattern
1	7
4[§]	3, 5, 6
5[§]	3, 5, 6
6	3, 4
7[§]	5
8[§]	5
9[*]	3, 8
13[*]	5
14[*]	5

[*] Subject 9 was suspected to have CMO II deficiency, while the brothers, subjects 13 and 14, were believed to have a defect in aldosterone synthase activity. All other subjects are presumed to have functional *CYP11B2* enzyme.

[§] Subjects 4 and 5 are related and gave patterns unique to their kindred. Similarly subjects 7 and 8 are brothers and gave a pattern unique to their kindred.

5- Characterization of mutations in *CYP11B1* and *CYP11B2*

5.1 Sequencing *CYP11B* DNA

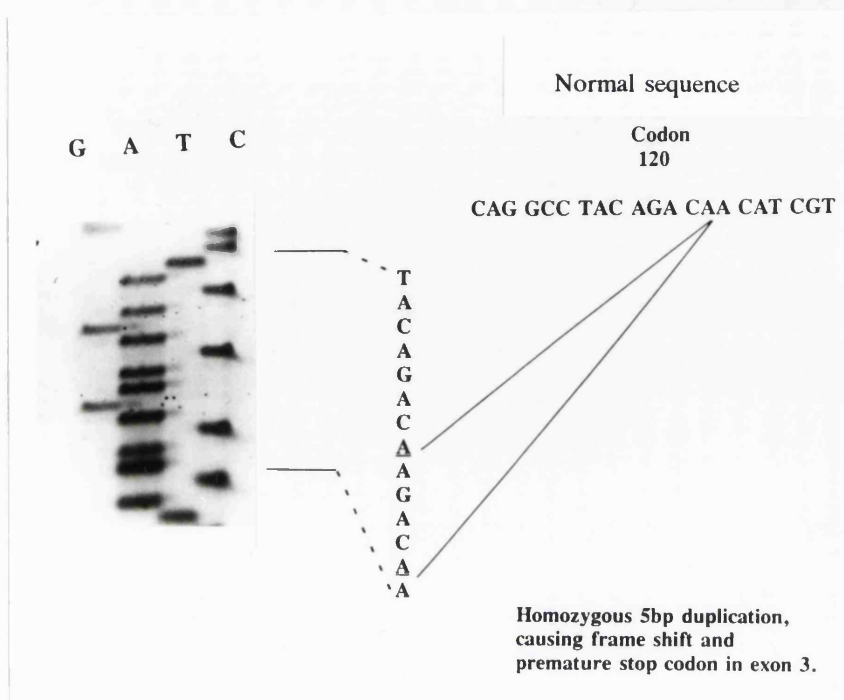
The regions of *CYP11B* DNA believed to contain unique sequence changes were sequenced directly from PCR amplified DNA or from cloned DNA, using the methods described in section 2.2.10. Cloning of PCR products in pUC18 plasmid was carried out as described in section 2.2.9. Sequencing cloned DNA was the preferred method, giving clearer results, especially for heterozygous individuals. Direct sequencing gave double bands at points of allelic difference in heterozygotes, but on occasion also gave non-specific double bands. The problems that arose from sequencing are discussed further in the evaluation of methods section of chapter 6.

5.2 Mutations in the *CYP11B* genes

5.2.1 Small duplication in *CYP11B1* exon 2 causing frame shift

PCR amplified DNA from subject 1 showed a markedly different SSCP pattern in exon 2 (Figure 4.6d). DNA amplified from a second initial PCR reaction was cloned into pUC18 and sequenced. As a third level of control, DNA was amplified once more from genomic DNA and used for direct sequencing (Winship, 1989). In all cases comparison with the published sequence (Mornet *et al*, 1989) showed a 5 bp duplication (AGACA) between codon 121 (AGA) and 122 (CAA) (Figure 5.2a), which would lead to a shift in reading frame, resulting in a premature termination codon at the beginning of exon 3. DNA from the parents was not available to confirm homozygosity of this mutation, but no normal sequence was detected at this point and polyacrylamide gel electrophoresis of exon 2 DNA from this individual

Figure 5.2a: Nucleotide sequence of *CYP11B1* exon 2 from subject 1, showing 5 bp duplication.



showed no normal size band. Hemizygosity was excluded by Southern blot analysis, which showed no difference in intensity of *CYP11B1* relative to *CYP11B2*. The homozygous presence of this mutation is the probable cause of 11ß-hydroxylase deficiency in this patient (Skinner & Rumsby, 1994). The process by which this 5 bp duplication may have occurred is discussed in section 6.2.1.

5.2.2 Small deletion in *CYP11B1* exon 2 causing frame shift

Following SSCP analysis, subject 6 was suspected to be heterozygous for a deletion in *CYP11B1* exon 2 (Figure 4.6d). This was confirmed on sequence analysis of cloned PCR amplified DNA, which showed the presence of a 28 bp deletion on one allele of exon 2 between codons 104 and 114 (Figure 5.2b). The sequence was obtained from a single clone, but direct sequencing showed that double bands consistent with the superimposition of normal and deleted allele sequences began at the expected point of deletion. A second clone, obtained from the same PCR amplified DNA, had a normal sequence at this point. Interestingly both clones also had a second mutation in intron 1, just 5' of exon 2, which occurred within a sequence required for intron splicing. This second mutation is discussed in section 5.2.5. The 28 bp deletion leads to an altered reading frame and a premature stop codon in exon 3. As the deletion was present on only one allele a second mutation was expected to be present and was subsequently found in exon 4 (section 5.2.4).

5.2.3 Nonsense mutation

A single homozygous nonsense mutation CAG→TAG, Q356X, was discovered in subject 4, following direct sequencing (Figure 5.2c). This mutation in exon 6 would produce a truncated protein without haem or substrate binding ability. The same mutation was reported by Curnow *et al* (1993), in a patient of African-American origin. Subject 4 was of Nigerian origin, which may suggest ancestral linkage with the American patient. The SSCP difference in exon 6 that targetted this region for sequencing was far more obvious in the father of the subject (Figure 4.6f) and indicates the benefit of analyzing family members together with affected individuals.

Figure 5.2b: Nucleotide sequence of *CYP11B1* exon 2 from subject 6, showing 28 bp deletion.

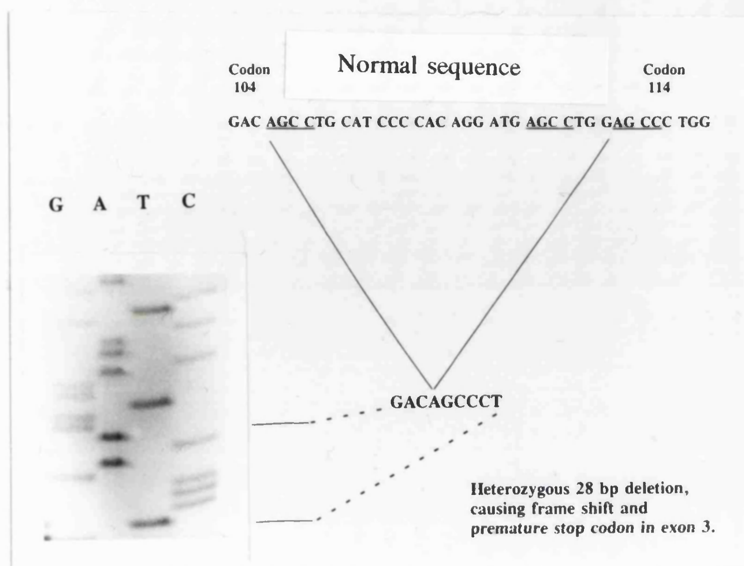
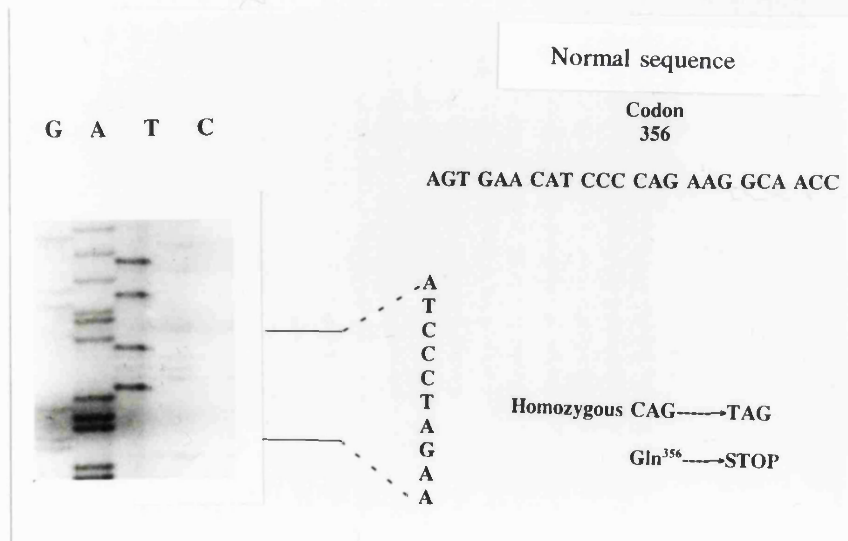


Figure 5.2c: Nucleotide sequence of *CYP11B1* exon 6 from subject 4, showing Q356X mutation.



5.2.4 Mis-sense mutations leading to amino acid changes

Subjects 1 and 6 both had unique mis-sense mutations in *CYP11B1* in addition to the lesions described in sections 5.2.1 and 5.2.2. The change in exon 8 of subject 1 was shown to be due to the sequence changes CGC→CAC, giving the amino acid change R427H, and GTA→GCA, V458A, as shown in Figure 5.2d. The R427H mutation removes a *Hha*I restriction enzyme site. Digestion of PCR amplified DNA from subject 1 with *Hha*I endonuclease and electrophoresis on a 10% polyacrylamide gel, alongside DNA with normal sequence and ϕX174 *Hae*III markers, confirmed the homozygous presence of this mutation (Figure 5.2e). The frame shift in exon 2 would make the changes in exon 8 irrelevant to the patients condition, but it is interesting to consider whether they arose before the mutation in exon 2 and whether they would affect enzyme activity if present in isolation.

The SSCP anomaly in *CYP11B1* exon 4 of subject 6 was found to be due to the homozygous sequence change GGC→CGC (G267R) (Figure 5.2f). This sequence change was found in four clones and in directly sequenced DNA from a separate PCR reaction. The amino acid change resulting from this mutation could be responsible for the loss of CYP11B1 function in this patient, although the patient was also homozygous for a mutation in intron 1 and heterozygous for the 28 bp deletion in exon 2, as discussed in section 5.2.2. The G267R mutation carries added significance, as it occurs within the left hand splice junction recognition site and may also affect intron splicing, as discussed in section 5.2.5.

The *CYP11B1* DNA from subjects 7 and 8 was also found to have a mutation that altered codon 267. In this case, however, the change was at the 5' end of exon 5 and involved the mutation GGC→GAC (G267D) (Figure 5.2g). The change was detected by direct sequencing and appeared to be on only one allele. Three attempts at cloning exons 4 and 5 as a single fragment and one attempt at cloning exon 5 DNA from subjects 7 and 8 failed to produce recombinant clones. More sequencing is required to confirm this mutation. It is significant, however, that this mutation affects the same amino acid as the mutation above, which could suggest that the glycine involved is important for enzyme structure and function. The presence of a second

Figure 5.2d: Nucleotide sequence of *CYP11B1* exon 8 from subject 1, showing R427H and V458A mutations.

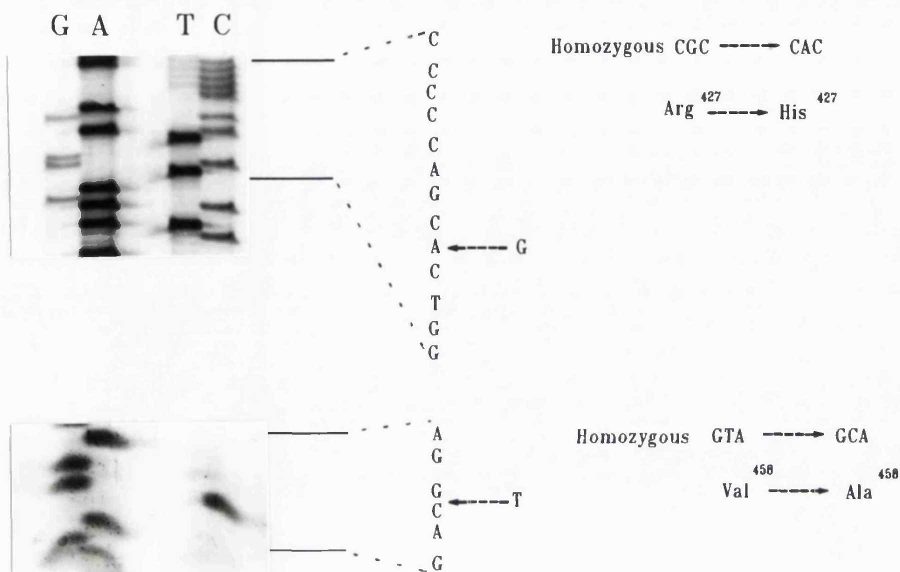
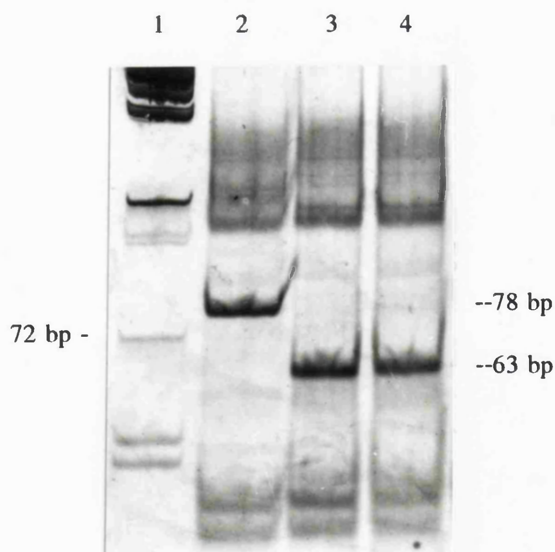


Figure 5.2e: *HhaI* restriction digest of exon 8 DNA from normals and subject 1, confirming R427H mutation.



Lane 1: ϕX174 *Hae*III size markers; lane 2: subject 1 DNA; lanes 3+4: normal control exon 8 DNA. The CGC→CAC mutation (R427H), removes a *HhaI* restriction site (GCGC). The normal digestion products of the 179 bp PCR product are 15, 18 (not shown on gel), 40, 43 and 63 bp. The loss of the *HhaI* site produces a digestion product of 78 bp, as shown. No normal sized band was produced from subject 1 DNA, which suggests that the R427H mutation is present on both alleles.

Figure 5.2f: Nucleotide sequence of *CYP11B1* exon 4 from subject 6, showing G267R mutation.

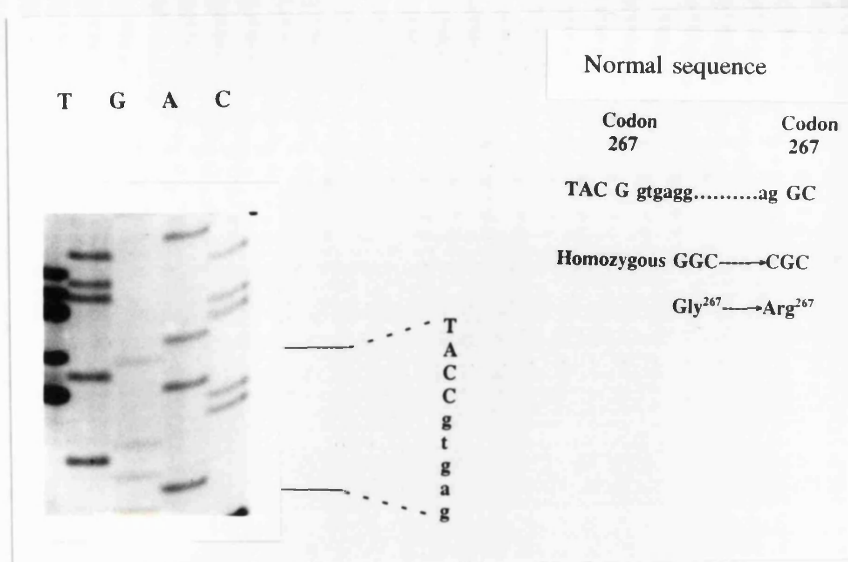
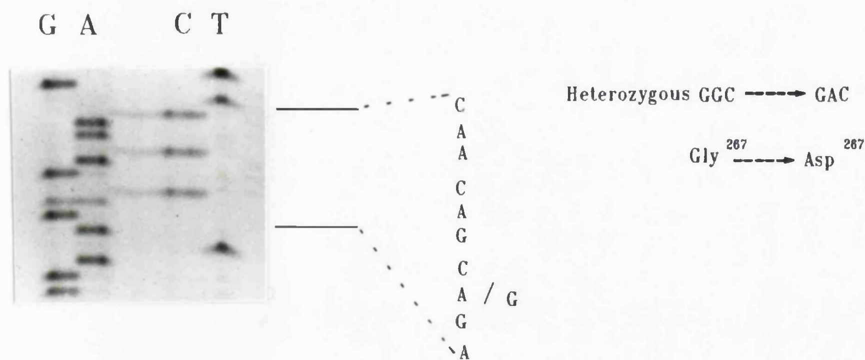


Figure 5.2g: Nucleotide sequence of *CYP11B1* exon 5 from subjects 7 and 8, showing G267D mutation.



The heterozygous presence of a mutation is shown by a G (normal) and A (mutant) band occurring at the same point in the sequence.

heterozygous mutation within these two subjects was expected at the 3' end of exon 5 from the SSCP pattern produced from *Rsa*I digested exon 4 and 5 DNA (Figure 4.6c). A mutation producing a similar SSCP pattern has been located in subjects 13 and 14 and is due to a sequence change that could affect intron splicing, as described in section 5.2.5. No sequencing data on the 3' region of exon 5 in subjects 7 and 8 is currently available.

A sequence change TGC→TTC, C494F, was found in exon 9 of both *CYP11B1* alleles from subjects 13 and 14 (Figure 5.2h). This mutation was also found in an unrelated individual with 11ß-hydroxylase deficiency, subject 10, and again appeared homozygous on direct sequencing, although this has not yet been confirmed. All three subjects carrying this mutation gave SSCP anomalies in exon 9, but the patterns of subjects 13 and 10 were not identical. This suggested that other sequence changes were present in these individuals in the region of exon 9. Additional sequencing showed that intron 8 differed between the two kindreds, but the changes did not occur in areas that would be expected to affect enzyme activity.

A final amino acid change was detected in the *CYP11B2* gene of subjects 13 and 14. This involved the change CGC→CAC, R470H and was detected by sequencing cloned DNA (Figure 5.2i). The homozygous nature of the mutation was confirmed by direct sequencing in both 13 and 14. This change is of interest in that CYP11B1 has histidine at position 470 and therefore this mutation could represent a gene conversion event. As there are CYP11B2 specific sequences in codons 469 and 471, which have not been altered, this explanation is unlikely. The mutation occurs at a CpG dinucleotide and is more likely the result of a single base change, as discussed in section 6.2.3. The SSCP pattern difference that targetted this region for further study in these subjects is most likely caused by an intronic change, AGGCCTGAT→ AGGCC$\underline{A}$GAT, a gene conversion to *CYP11B1* sequence, which leads to the loss of a *Stu*I enzyme site in intron 8 just 5' of the beginning of exon 9. As *Stu*I was used to digest SSCP samples for this region, the SSCP polymorphism in exon 9 of *CYP11B2* would be due to difference in digestion product size (178, 210 and 15 bp as opposed to 78, 100, 210 and 15 bp for the normal sequence).

Figure 5.2h: Nucleotide sequence of *CYP11B1* exon 9 from subjects 13 and 14, showing C494F mutation.

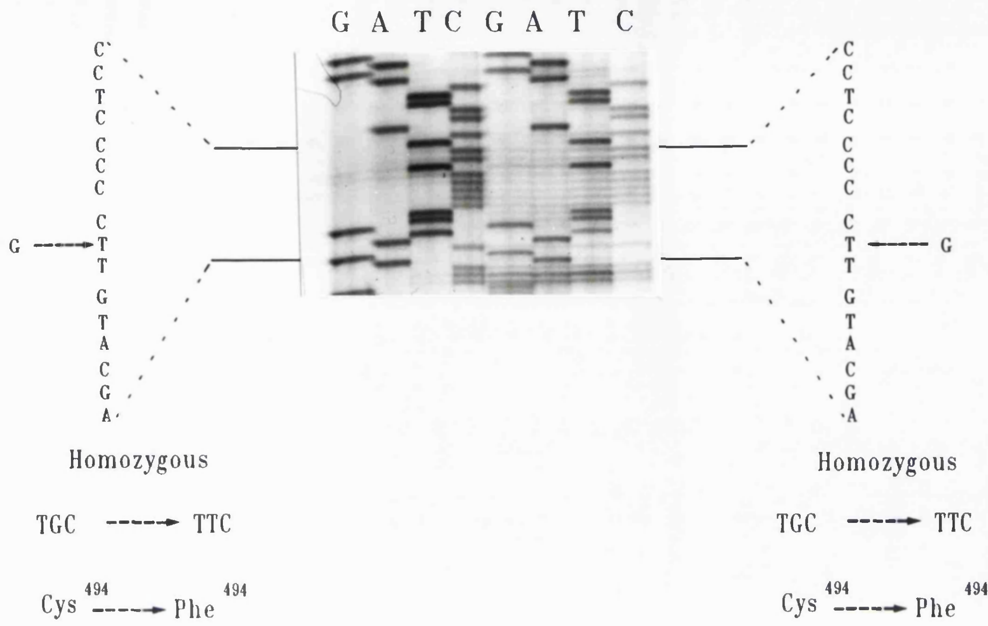
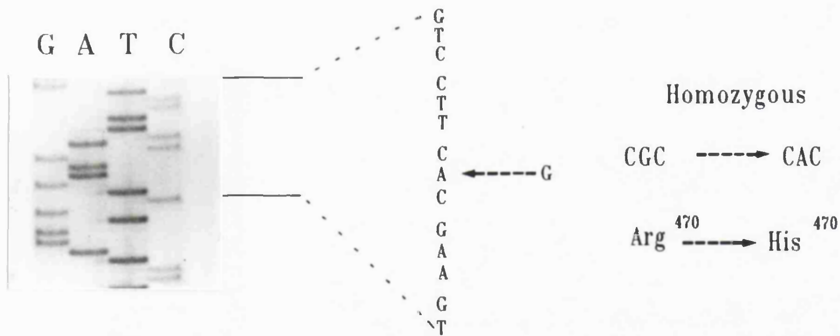


Figure 5.2i: Nucleotide sequence of *CYP11B2* exon 9 from subject 14, showing R470H mutation.



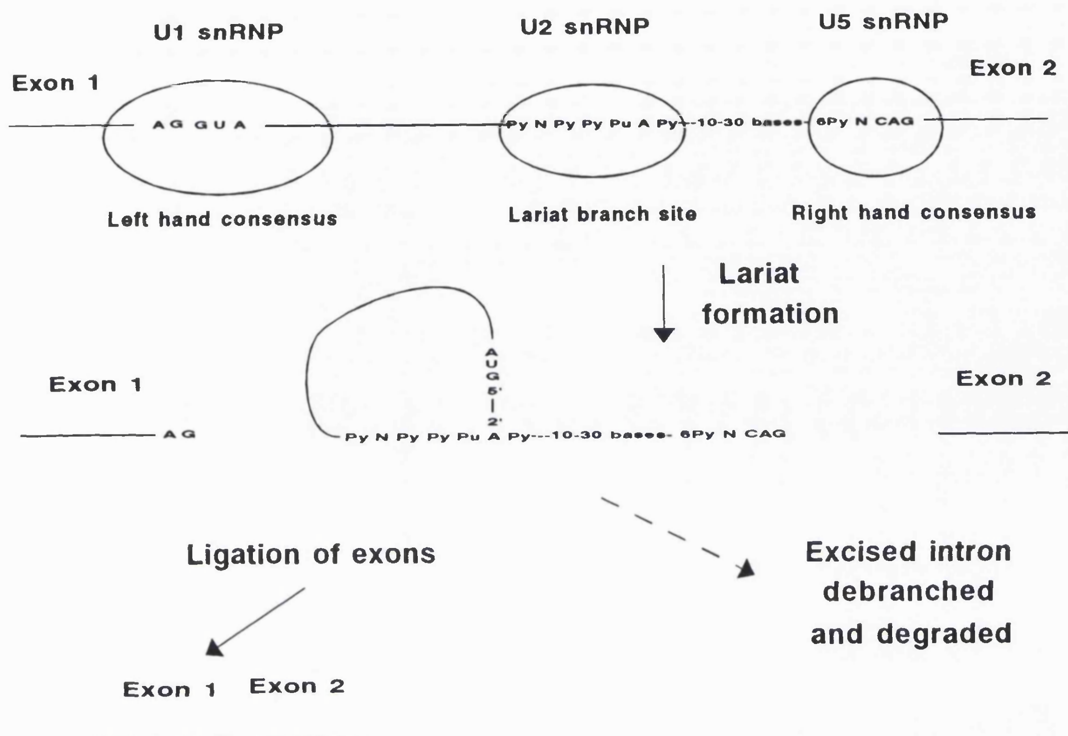
5.2.5 Intron splicing mutations

The mutations described in the preceding sections are all likely to have a significant effect on enzyme function. Other mutations were found that may have less effect on the production of functional enzyme. Three mutations within intron splicing regions fall into this category. Correct splicing of eukaryote nuclear encoded mRNA involves left and right hand intron recognition sites and a lariat branch site (Figure 5.2j). The most important markers of intron boundaries are the GT and AG dinucleotides, which define the 5' and 3' ends of the majority of introns. The left and right hand boundaries exist within consensus sequences, however, which are also important. Mutation of the G nucleotide immediately 5' of the GT dinucleotide of the left hand splice junction has been shown to lead to alternative splicing (Huang *et al*, 1993). Whilst not abolishing the production of correctly processed mRNA, such mutations can lead to a reduction in the level of mRNA in the cytoplasm and reduce functional protein present.

Two such mutations were found in subjects with deficiency of CYP11B1 enzyme. The G267R mutation described for subject 6 occurred at the 3' end of exon 4, altering the G nucleotide, normally found immediately 5' of the left hand splice junction of intron 4, to a C (Figure 5.2k). In subjects 13 and 14 the corresponding G nucleotide in exon 5, immediately 5' of the left hand splice junction of intron 5, was found to be mutated to an A (Figure 5.2l). SSCP analysis inferred that the same mutation may also be present, in the heterozygous form, in *CYP11B1* exon 5 DNA of subjects 7 and 8. In all cases these mutations were combined with mutations causing amino acid changes.

Subject 6, already found to be affected by a heterozygous deletion and a homozygous mutation leading to an amino acid change, was also found to carry a homozygous deletion of five base pairs within intron 1, with subsequent loss of the lariat branch site. The deletion was not a straightforward loss of five bases. Preceding the deletion was a four base pair duplication and a small gene conversion event, as shown in Figure 5.2m. The same sequence was obtained from two clones obtained from the same PCR amplified DNA.

Figure 5.2j: Intron splicing of most nuclear mRNAs requires three consensus sequences and the binding of proteins that contain small nuclear RNAs (snRNAs).



Intron recognition occurs when a protein containing U1 snRNA (U1 snRNP) binds to the left hand splice junction consensus sequence of mRNA (Maniatis & Reed, 1987). The interaction requires base complementarity between the left hand junction and the U1 RNA. A spliceosome forms, with proteins being recruited to the site of the intron. Within the spliceosome is a protein containing U2 snRNA (U2 snRNP). This has a degree of complementarity with a sequence that forms the lariat branch site, and binds at this site 10-30 bases 5' of the right hand consensus sequence. When the spliceosome is formed cleavage occurs at the left hand junction. The invariant G residue at the left hand junction is covalently joined to the A residue within the lariat branch site, by a 5'-2' phosphodiester bond, to form a lariat. The right hand splice junction is recognised by interaction of the junction consensus with a protein containing U5 snRNA (U5 snRNP). Cleavage of the right hand junction and ligation of the exons completes the splicing process. The excised intron lariat is debranched and degraded (Wassarman & Steitz, 1993).

Figure 5.2k: The G267R mutation in subject 6 occurs within the left hand splice junction of intron 4 and reduces the complementarity with U1 snRNA.

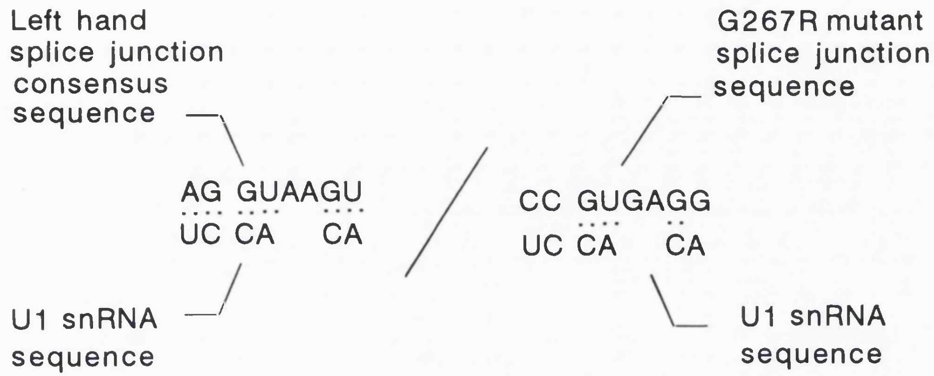


Figure 5.2l: Nucleotide sequence of *CYP11B1* exon 5 of subjects 13 and 14 contains a G→A change that occurs at the left hand splice junction of intron 5.

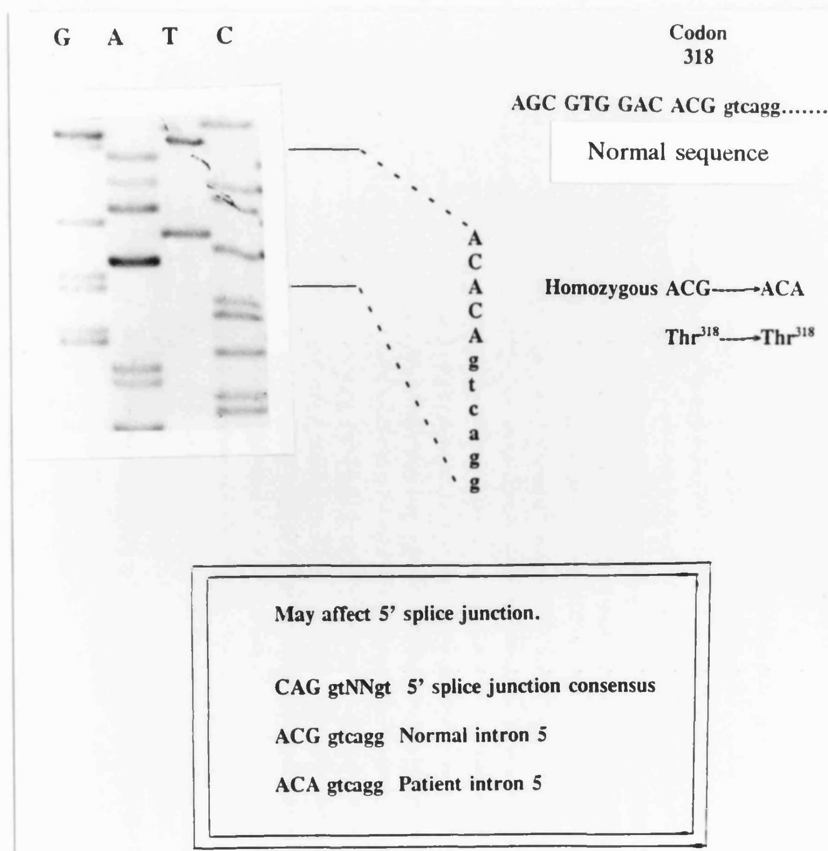


Figure 5.2m: Nucleotide sequence of *CYP11B1* intron 1 from subjects 6 and 10, showing loss of intron lariat branch site.

G A T C

Subject 6	Normal	Subject 10
TAG CAT G gacaccccggttcgtgctccggcacgacctacctggtacaggg	TAG CAT G gacaccccggttcgtcctccgggacgacctgccccagatcaggg	CAG CAT G gacaccccggttcgtgctccggtacgacctacctggtacaggg + (C) + (t)

G A T C

The above diagram shows the normal sequence for *CYP11B1* 3'intron 1, and the sequence for the same region in subjects 6 and 10. In both of these subjects 5 bp (acccc) have been deleted. This removes the adenosine nucleotide required as part of the lariat branch site, the consensus for which is Py N Py Py Pu A Py. In addition to the deletion, there is a 2 bp gene conversion, ta→at, 4 bp duplication (tcca), and several point mutations as indicated. The two subjects differ at the two points marked by +.

Direct sequence analysis of subject 10 showed an identical mutation, again in the homozygous form. A cryptic branch site exists just 5' of the region affected, but consists of the sequence GAGGGAC, which conforms poorly with the consensus branch sequence Py N Py Py Pu A Py, where Py= pyrimidine and Pu= purine, e.g. TACTAAC. Mutations leading to loss of lariat branch sites do not usually lead to total loss of splicing in higher eukaryotes, as cryptic sites containing A nucleotides are often located nearby. These cryptic sites are not, however, recognised as efficiently by the intron processing enzymes (Lewin, 1990).

SSCP analysis of exon 2 targetted this region as carrying a sequence anomaly in both subjects 6 and 10, but the anomalous patterns were quite different. One explanation for this is that the sequence of subject 10 differed from that of subject 6 at two points, one within intron 1 (as shown in Figure 5.2m) and the other in codon 82, where subject 10 had the silent mutation GAT→GAC. The SSCP difference observed in subject 10 may well represent these two base changes rather than the intron 1 lesion, which occurs near the 5' end of the amplified fragment.

The presence of the intron 1 mutation within two unrelated individuals, and the absence of a normal sequence at this point in either subject, would suggest that this may well be a rare polymorphism, which could reduce the level of correct splicing. In the case of subject 6 the contribution of the two other mutations would be the presumed cause of loss of *CYP11B1* function. In subject 10, the intron 1 mutation occurs in combination with the amino acid change C494F, which in turn could lead to 11ß-hydroxylase deficiency if splicing efficiency was reduced.

<u>5.2.6</u> <u>Silent exonic mutations</u>

Three silent exonic mutations were discovered following sequencing of regions targetted by SSCP analysis. In exon 1 of subjects 13 and 14, the change CTG→CTA, L75L was found on sequencing cloned DNA from both brothers. SSCP analysis suggested the presence of only 2 variants of exon 1. This being so, the above silent mutation could represent the common variant shown by SSCP analysis.

A silent mutation was also found in exon 2, as discussed above, involving the sequence change GAT→GAC, D82D. This was found in subjects 1, 10, 13 and 14 and again may represent a common polymorphism. A subtle decrease in mobility was observed on SSCP analysis of exon 2 DNA from individuals homozygous for this change.

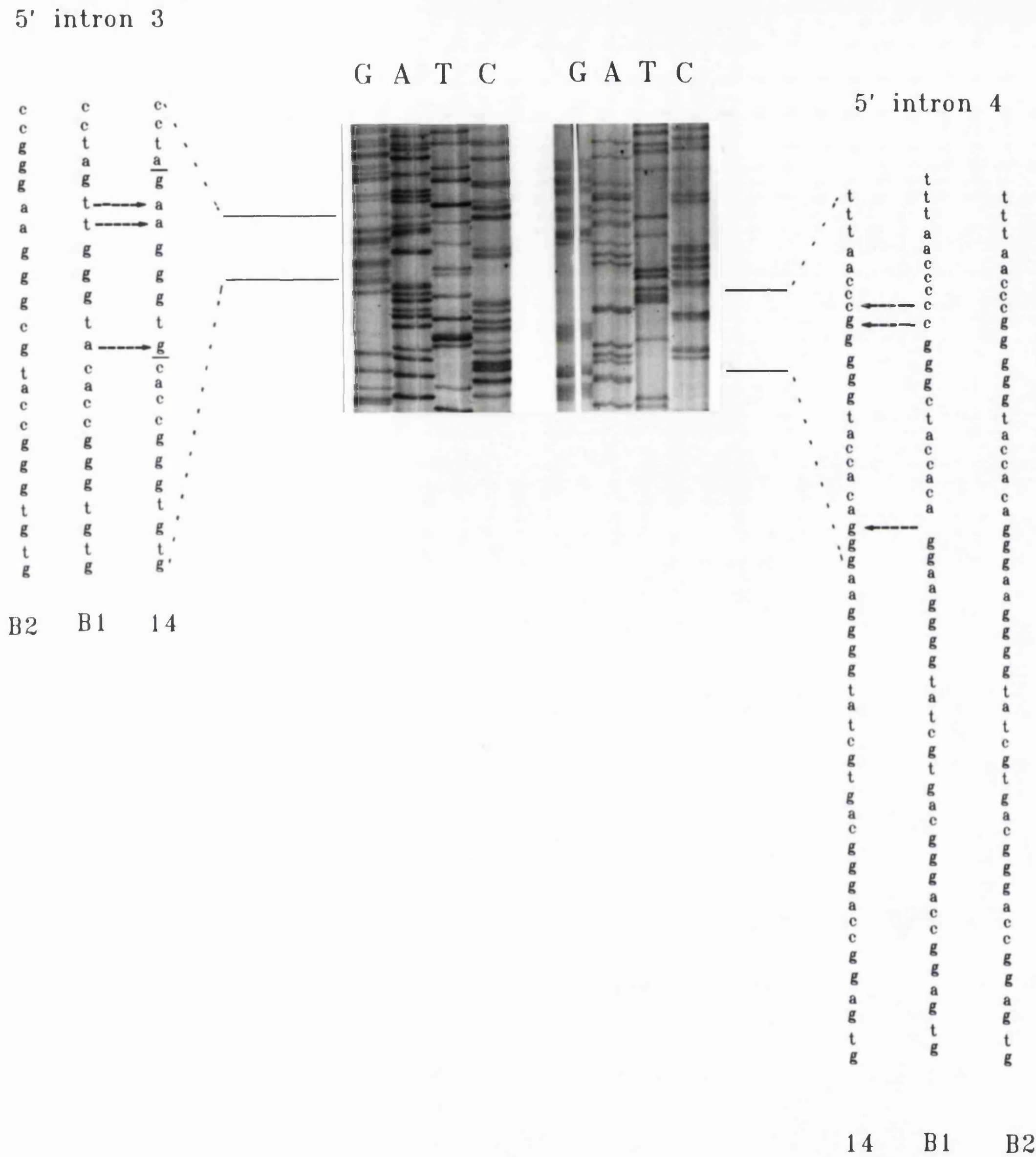
A silent mutation is believed to be the cause of the unique SSCP pattern of subject 1 in the region of exon 6. In this case the change is GCA→GCC, A348A. The SSCP difference was very pronounced and this illustrates the fact that not all SSCP differences are caused by deleterious or particularly significant mutations.

<u>5.2.7</u> <u>Gene conversions in *CYP11B1*</u>

One of the aims of the study was to determine whether gene conversion and non-homologous recombination lead to increased mutation at the *CYP11B1* locus. None of the deleterious mutations listed above represented gene conversions or non-homologous recombination between the two *CYP11B* genes.

Points of intronic gene conversion were located, however, in both *CYP11B1* and *CYP11B2*. Five regions of apparent gene conversion were located in the *CYP11B1* gene of subjects 13 and 14, during the course of sequencing cloned DNA, two of which are shown in Figure 5.2n. Three of the apparent gene conversion events had *CYP11B1* specific sequence at each end, which defined the limits of the region of gene conversion. In intron 2, there were two C→T changes 30 bases apart, which would alter the region to the sequence in *CYP11B2*. *CYP11B1* specific sequence

Figure 5.2n: Nucleotide sequences of 2 regions of gene conversion from *CYP11B1* to *CYP11B2*



In the above diagram the nucleotide sequences of 5' intron 3 and 5' intron 4 are shown (B2= published sequence of *CYP11B2*, B1= published sequence of *CYP11B1*, 14= *CYP11B1* sequence obtained from subject 14). In intron 3 of subject 14, an 8 bp region had undergone apparent gene conversion from *CYP11B1* sequence to that in *CYP11B2*. In intron 4 a region from 92-182 bp had been converted to *CYP11B2* sequence. In each case *CYP11B1* specific sequences defined the limits of gene conversion and the extent of conversion was given by base differences at either end of the converted region. The published sequences are shown in Appendix 4 and 5. These changes could also represent an ancestral form of the *CYP11B1* gene that bore more resemblance to *CYP11B2* (see text for details).

occurred 15 bases to either end of this region, which meant that the length of the conversion in this case could be anything from 32-60 bp.

In intron 3 the sequence was altered from ATGGGTTG→GTGGGAAG, with *CYP11B1* specific sequence to either side, implying that only 8 bp were converted. This may well represent the common polymorphism present in the exon 3 region of *CYP11B1*, as described earlier in section 4.6.2. Direct sequencing of subject 3, suspected to be heterozygous for this polymorphism, gave a mixed sequence at this point as would be expected. In intron 4 the following change was found: GGACACCATCGGGCCCCC → GGGACACCATGGGGGCCC. Here the length of conversion could be anything from 92-182 bp. This gene conversion may represent the common polymorphism in the exon 4 region.

The length of two other conversions cannot be given with any certainty. At the 5' end of intron 6 the change was from AACCCCATCCCAGCTGAGA →AAGCCCCATCCAGCTGAGGA, so that the gene conversion is greater than 16 bp in length. The last gene conversion was at the 3' end of intron 6 and involved the change CCCTCGAG→CCCCTCGAG.

It would seem likely that gene conversions vary from around 10 to 100 bp. The change at the 3' end of intron 6 is interesting in that it forms a sequence similar to the right hand splice junction, with the remote chance that it could lead to altered splicing. However, the fact that *CYP11B2* has this sequence, but does not appear to use it as a right hand splice site, would suggest that this change does not affect splicing.

Gene conversion events seemed to be most common in subjects 13 and 14, but this may reflect the fact that the majority of the gene was sequenced for these two individuals, whereas the genes of other individuals were only partially sequenced. Aside from the possible gene conversion in intron 1 of subjects 6 and 10 (see section 5.2.5), another possible point of gene conversion was found in intron 8 of subject 10 (CCACGTCGATGGGCT→CCACGTGCATGGGCT). It would seem that gene conversion does alter *CYP11B1* intron sequences and that these differences could be useful as polymorphic markers.

5.2.8 Gene conversions in *CYP11B2*

Sequence analysis of exons 5 and 9 of the *CYP11B2* gene from subjects 13 and 14 revealed four regions of gene conversion. Two small gene conversion changes were shown to occur at the 3' end of intron 4: one near to a putative branch site for lariat formation, the other at the right hand splice junction (Figure 5.2o). The first gene conversion, involving up to 11 bp, introduces a second branch site just 3' of the usual branch site consensus (Figure 5.2p). This arrangement is present in *CYP11B1* and probably does not affect splicing. The gene conversion also inserts a *Dde*I restriction enzyme site. On restriction digest analysis of both affected subjects, the parents and seven unaffected siblings all were shown to be homozygous for this gene conversion (Figure 5.2q).

The second gene conversion, at the right hand splice junction, involves only 3 bp and inserts a G into the pyrimidine core of the splice site consensus. This change removes an *Rsa*I restriction enzyme recognition site. Again, digestion of DNA showed all family members to be homozygous for this change (Figure 5.2r). This would suggest that these intron changes do not affect splicing. Individuals from other kindreds were not tested for the above sequence changes.

SSCP analysis had shown the two affected brothers to have a unique pattern for exon 5 of *CYP11B2*. This was shown to be due to a change involving a single missing C (CAAG<u>C</u>CTGCC→CAAGCTCC) in the intron 44 bp 3' of exon 5, which was not a gene conversion and is presumed to have no effect on splicing.

A third gene conversion was found in intron 8 of subjects 13 and 14, just 5' of exon 9. This was shown by only a single nucleotide change GCCTGATG→GCC<u>A</u>GATG, but may represent a gene conversion event spanning up to 12 bp. This removes a *Stu*I restriction enzyme recognition site and would account for the SSCP polymorphism shown in exon 9 of the *CYP11B2*, which was produced by *Stu*I digestion of a radioactive 403 bp fragment including exon 9. *Stu*I digestion of exon 9 DNA from subjects 13 and 14 and their parents showed that the father and sons were homozygous for this polymorphism, while the mother was heterozygous. As such it cannot be used as a marker for tracing inheritance of the defective gene in this family. The change occurs within a branch site consensus region, but presumably does not affect splicing as it is present on both alleles of unaffected individuals.

Figure 5.2o: Nucleotide sequence showing *CYP11B2* to *CYP11B1* gene conversion in intron 4 of *CYP11B2*.

G A T C

T G a c a g g t c c c t c c c c c t g t c g g g a g t c

Normal sequence

T G a c a t g t c c c t c c c c c t g t c g g g a t c

Figure 5.2p: Gene conversion introduces a second lariat branch site for intron splicing.

Py N Py Py Pu A Py = consensus branch site

caccaag atctgag

cttcccagcaccaagatct g agggctgtcccctgctccctg g acag GT **Subject 13 and 14 intron 4**

The mutation in *CYP11B2* intron 4 of subjects 13 and 14 introduces a second lariat branch site, atctgag, immediately after the normal branch site for *CYP11B2* intron 4, caccaag. The same arrangement is found in *CYP11B1*.

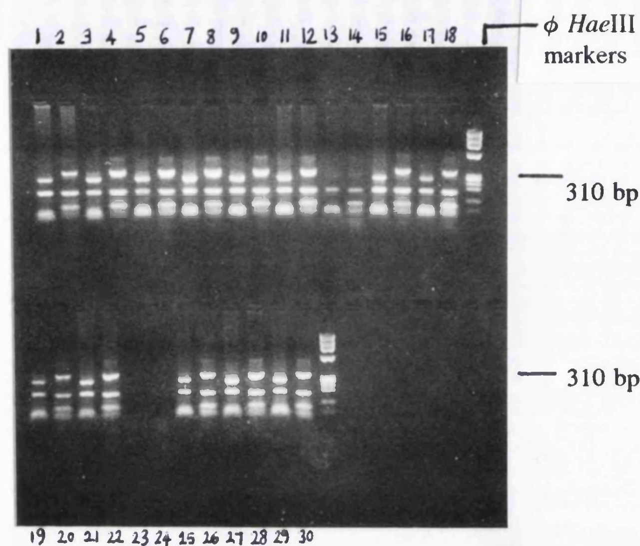
Figure 5.2q key:

Lanes 1+2: digested and undigested samples of father
3+4: " " " " " mother
5+6: " " " " " subject 13
7+8: " " " " " subject 14
9-30: " " " " " seven siblings.

Figure 5.2r key:

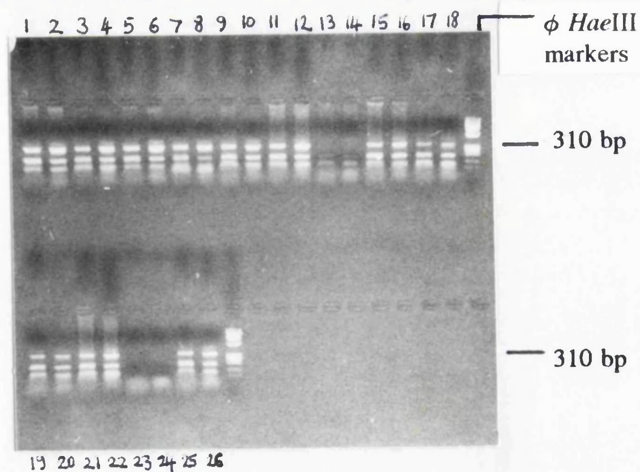
Lanes 1+2: undigested and digested samples of father
3+4: " " " " " mother
5+6: " " " " " subject 13
7+8: " " " " " subject 14
9-26: " " " " " seven siblings.

Figure 5.2q: *Dde*I restriction digest of *CYP11B2* intron 4/ exon 5 DNA from the family of subjects 13 and 14, to confirm the presence of a gene conversion at the intron lariat branch site.



The apparent *CYP11B2*→*CYP11B1* gene conversion (..ctaggg→..ctgaggg) inserts a *Dde*I restriction site (CTNAG). No *Dde*I site is present in this region in the published sequence of *CYP11B2*. Digestion of the 308 bp PCR product containing intron 4/ exon 5 DNA yielded the expected fragment of 268 bp, for the two affected brothers, their parents and seven unaffected siblings. Above is shown an ethidium bromide stained agarose gel, with digested and undigested DNA samples for each of the family members. These results suggest that the gene conversion event is present in all family members in the homozygous form.

Figure 5.2r: *Rsa*I restriction digest of *CYP11B2* intron 4/ exon 5 DNA from the family of subjects 13 and 14 to confirm the presence of a gene conversion in the right hand splice consensus region.



The apparent *CYP11B2*→*CYP11B1* gene conversion (..gtacag GT→..ggacag GT) removes an *Rsa*I recognition site (GTAC). The expected digestion products (246 and 62 bp) of the 308 bp PCR fragment containing intron 4/ exon 5 DNA were not observed, for the two affected brothers, their parents and seven unaffected siblings. Above is shown an ethidium bromide stained agarose gel, with undigested and digested DNA samples for each of the family members. The top band is the 308 bp fragment and showed no digestion. These results suggest that the gene conversion event is present in all family members in the homozygous form.

The final apparent gene conversion occurred in the 3' untranslated region of *CYP11B2* and may involve a region up to 43 bp in length. The only difference in this region is a single nucleotide CCCACGTGCACA→CCCACATGCACA. There was also a mutation in this region involving the insertion of a T. The effect of changes in this region are unknown. The polyadenylation sites required for 3' processing of mRNA occur downstream of this sequence.

Less sequencing work was carried out on the *CYP11B2* gene, but a fourth point of apparent gene conversion was found in intron 2 of subject 9, the patient with CMO II deficiency (GGGATACGGGGTCA→GGGATACGGGTCA). The possibility that a proportion of the above changes do not represent gene conversions is discussed in section 6.2.7.

A summary of significant mutations found in each subject is given in Table 5.1.

Table 5.1: Mutations characterized in affected subjects.

Subject	Mutations found in CYP11B1							
	Intron 1	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 8	Exon 9
1		5bp Dupl. 120-121				A348A*	R427H V458A	Intron change
3			a					
4						Q356X		
6	Loss of branch site	28bp Del. 104-114		G267R				
7					G267D+ Splice Mutation?			
8					G267D+ Splice Mutation?			
10	Loss of branch site							C494F
13		D82D*			Splice Mutation?			C494F
14		D82D*			Splice Mutation?			C494F
15				a				
16				a				

Subjects	Mutations found in CYP11B2		
	Exon 3	Exon 8	Exon 9
9	a	a	
13			R470H
14			R470H

a These regions were different on SSCP but have not yet been fully characterised (see section 5.3).

* These represent silent mutations GAT→GAC at codon 82 and GCA→GCC at codon 348.

5.3 Mutations requiring verification

SSCP analysis suggested that *CYP11B1* exon 4 of subjects 15 and 16, *CYP11B1* exon 3 of subject 3, and *CYP11B2* exons 3 and 8 of subject 9, all contained mutations. Characterization of mutations at these points proved difficult, with problems in cloning and obtaining clear sequences at key points of interest.

Direct sequencing of exon 4 DNA from subjects 15 and 16 suggested the presence of a heterozygous mutation ACC→ATC, T248I. Such a mutation could represent a gene conversion from *CYP11B1* to *CYP11B2*.

Sequencing of *CYP11B2* DNA from subject 9 indicated a sequence change CGG→TGG, R181W on one allele. A homozygous change of R181W has been implicated in several cases of CMO II deficiency in subjects of Iranian origin.

Direct sequencing of *CYP11B1* exon 3 from subject 3 failed to detect a deleterious mutation, but revealed the presence of the same polymorphism in intron 3 as was found in subjects 13 and 14 (section 5.2.7). Further work is required to verify the above mutations and identify the cause of 11ß-hydroxylase deficiency in subject 3.

5.4 PCR artifacts and non-reproducible mutations

Analysis of subjects 13 and 14 was made more difficult by the presence of spurious mutations, which were detected only once and could not be reproduced on sequencing DNA from a second PCR reaction. Six of these were obtained in the course of sequencing the nine exons of *CYP11B1* for these two subjects. The existence of PCR and cloning artifacts is discussed in section 6.5.

6- Summary of results and general discussion

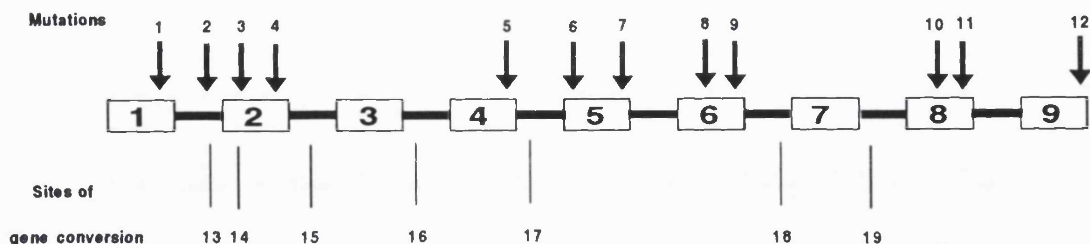
6.1 Mutations found in *CYP11B1*, *CYP11B2* and *CYP17* genes

All mutations located by this study, in both *CYP11B* genes and *CYP17*, are shown in Figures 6.1a and 6.1b. Eleven mutations causing loss of CYP11B1 function have been previously characterized. Of these, four involve premature stop codons W116X (Naiki *et al*, 1993), K174X, Q338X, Q356X (Curnow *et al*, 1993). Two involve frame shifts: a deletion of 1 bp from codon 32 (Curnow *et al*, 1993) and a duplication of 2 bp in codon 394 (Helmberg *et al*, 1992). The other five are amino acid changes: T318M, R374Q, R384Q, V441G (Curnow *et al*, 1993) and R448H (White *et al*, 1991).

This study has added two more additional frame shift mutations to this total: the 5 bp duplication at codon 120 and the 28 bp deletion at codon 104. Five additional amino acid changes G267R, G267D, R427H, V458A, C494F in the CYP11B1 have been characterized from DNA analysis of individuals affected by 11ß-hydroxylase deficiency. The effects of these amino acid substitutions are as yet unknown, but their discovery sets the stage for further work to determine how these changes affect enzyme function.

Changes have been found that may affect splicing of introns 1, 4 and 5. Such changes have not been previously reported in the *CYP11B1* gene. There is a possibility that these may represent rare polymorphisms, which reduce 11ß-hydroxylase production, but do not lead to disease unless combined with other, more deleterious mutations. Further studies to evaluate the effect of splice variants on 11ß-hydroxylase activity provides another avenue of research into adrenal related hypertension.

Figure 6.1a: Mutations located in *CYP11B1*



	Frame shift	Patients
3	28bp deletion of codons 105-113	6
4	5bp duplication AGACA at codon 121	1

	Intron splicing	Patients
2	actagaccccgtcca-> acatggtccatcca	6 and 10
5	ACG gtgagg-> ACC gtgagg	6
7	ACG gtcagg-> ACA gtcagg	13 and 14

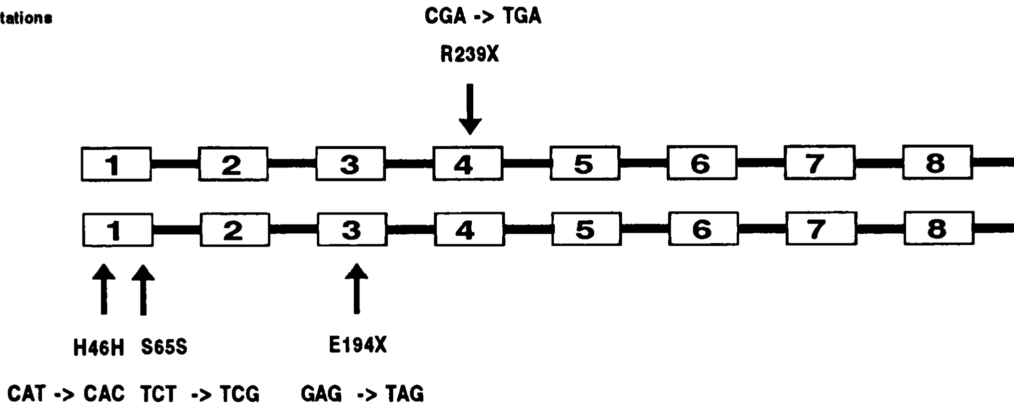
	Gene conversions	Patients	Converted region
13	ctag->catg	6 and 10	—
14	D82D GAT->GAC	10, 13 and 14	—
15	c->t - - - - - c->t	13 and 14	60bp
16	atgggttg->gtgggaag	" " "	8bp
17	cgggccccc->cgggggccc	" " "	100bp
18	aaccccatcccagctgaga-> aagccccatccagctgagga	" " "	—
19	ccctcgag->cccctcgag	" " "	—

	Amino acid change	Patients
5	G267R GGC->CGC	6
6	G267D GGC->GAC	7 and 8
9	Q356X CAG->TAG	4
10	R427H CGC->CAC	1
11	V458A GTA->GCA	1
12	C494F TGC->TTC	10, 13 and 14

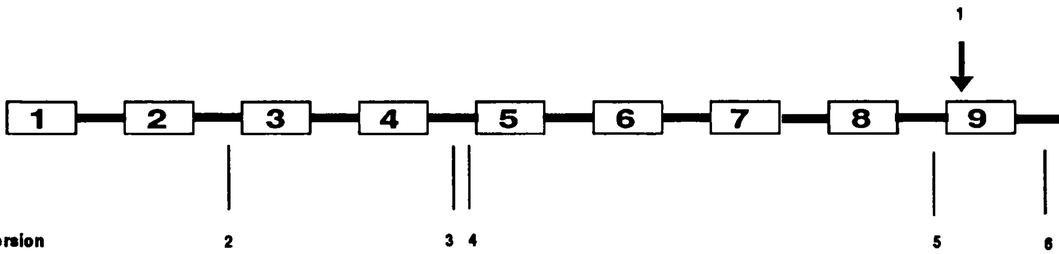
	Silent mutations	Patients
1	L75L CTG->CTA	13 and 14
8	A348A GCA->GCC	1

Figure 6.1b: Mutations located in the *CYP17* and *CYP11B2* genes.

CYP17



CYP11B2



	Amino acid change	Patients
1	R470H CGC->CAC	13 and 14

	Gene conversions	Patients	Converted region
2	tacgggg -> tacggg	9	—
3	caagatctaggg -> caagatctgaggg	13 and 14	7bp
4	ccctgtcag GT -> ccctggcag GT	" " "	3bp
5	aggcctgataaaccc -> aggccagataaaccc	" " "	15bp
6	cttctctcccacgtgc -> cttctctcccacatgc	" " "	43bp

The polymorphisms in exons 1, 3 and 4 of the *CYP11B1* gene, located during SSCP analysis, are presumed to be due to the sequence variations identified, all of which are unlikely to have any effect on enzyme function. The clear division shown by these polymorphic groups, and the simplicity of their detection by SSCP, makes them useful as markers in heredity studies of the role of the *CYP11B1* gene in disease. Similarly, the polymorphisms within *CYP11B2* exon 2 and 9 may prove to be useful markers. The fact that the subjects studied were clearly divided into two polymorphic groups, but that subjects 4 and 5 belonged to one group when *CYP11B1* was analyzed, but belonged to the other group when *CYP11B2* was examined, is also of interest and could suggest that *CYP11B1* and *CYP11B2* have diverged to different extents in different populations.

The work with *CYP17* and *CYP11B2* was less detailed, partly because of the availability of fewer patients with disorders of these genes, but also because of time constraints. Mutations were, however, found in both of these genes. In the *CYP17* two premature stop codons were found, one unreported and one previously reported in a Canadian subject (Ahlgren *et al*, 1992). Drawing conclusions on the rate of mutation within *CYP17* as compared to *CYP11B1* is difficult on the basis of the study carried out, as only a single *CYP17* patient was available. However, published reports, characterizing 16 deleterious mutations in *CYP17* suggest that there is far less variation in *CYP17* sequence when compared to the changes observed by this study of *CYP11B1*.

SSCP analysis of *CYP11B2* detected polymorphisms and found a single amino acid change R470H in two subjects affected by aldosterone synthase deficiency. To date only three mutations associated with defects of CYP11B2 enzyme have been reported: R181W, on the same allele as a V386A mutation (Pascoe *et al*, 1992b), and a recently reported gene conversion event changing exon 3 and 4 from *CYP11B2* to *CYP11B1* sequence (Fardella *et al*, 1994b).

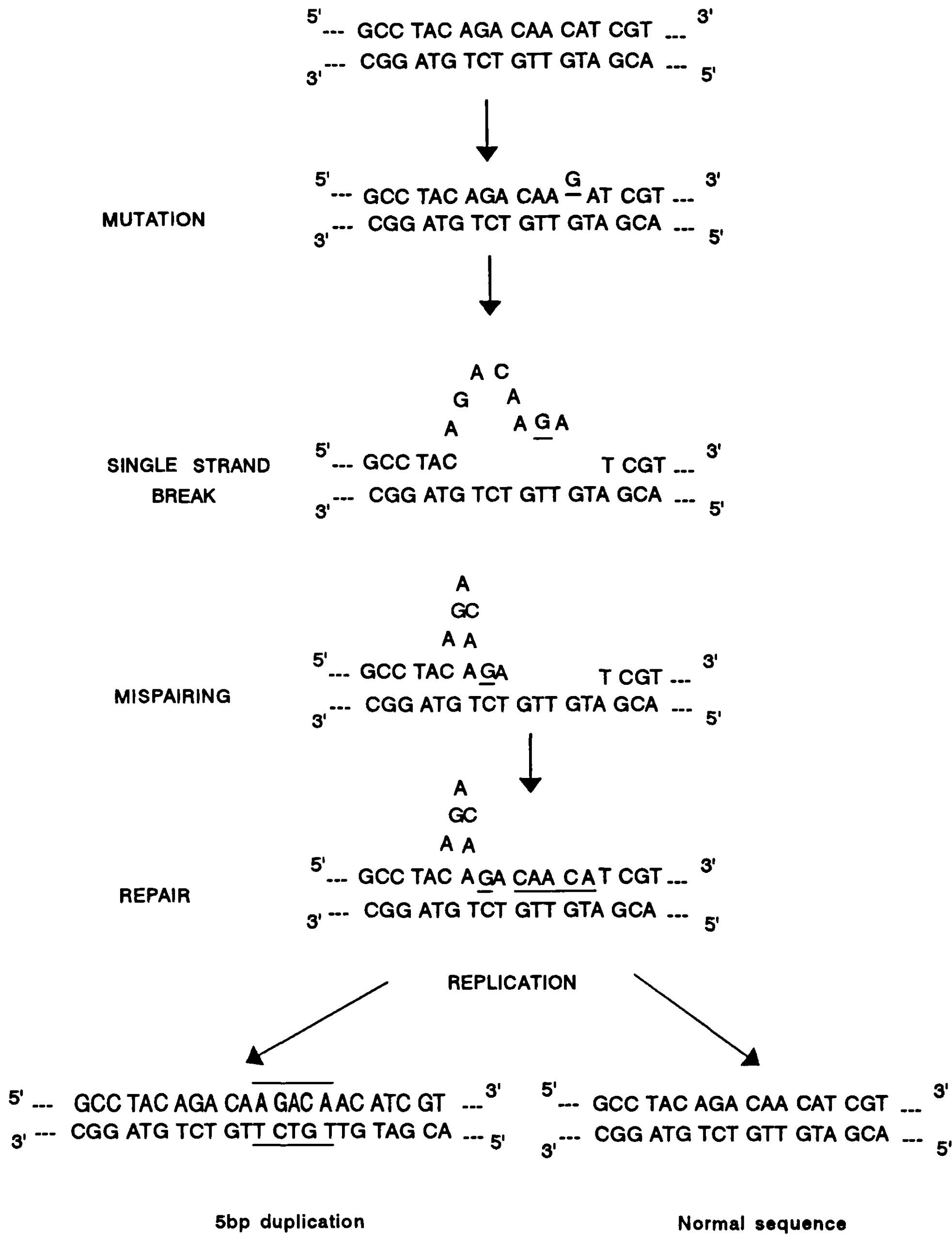
6.2 Mechanisms of mutation

6.2.1 Slipped strand mispairing

Slipped strand mispairing has been cited as a process involved in expanding repetitive DNA in introns, and a general cause of duplications and deletions in areas of repetitive DNA (Levinson & Gutman, 1987). This process is proposed to be the cause of the 5 bp duplication in exon 2 of subject 1 (Figure 6.2a). In this scenario a single mutation of a C to a G may have occurred at the beginning of codon 122, TACAGACAAC̲AT→TACAGACAAG̲AT, by random mutation or misalignment in this repetitive region and mismatch repair of DNA. With this mutation inserted, a single strand break occurs where the mismatch is generated. The 3' end of the broken strand mispairs with the TCT triplet of bases normally complementary to the AGA of the repetitive sequence shown above, leaving a 5 bp gap in the coding strand. This gap is repaired and the 5 base duplication has been introduced into the upper strand. When strand replication occurs prior to meiosis, an allele carrying a 5 bp duplication is produced.

This possible explanation is based on the repetitive nature of the DNA at codons 120 and 121. It is interesting that in *CYP11B2* codon 119 is not TAC, but ATC. This reduces the repetition in this region, removing an ACA repeat. This would reduce the chance of a similar mutation occurring in the *CYP11B2* gene. Such differences in the degree of repetition within exons may be partly responsible for the greater number of unique SSCP patterns in exon 2 of *CYP11B1* as compared to *CYP11B2*.

Figure 6.2a: Slipped strand mispairing may cause duplications in repetitive DNA



Shown above is the sequence of *CYP11B1* exon 2 at the point of the 5 bp duplication mutation located in the DNA of subject 1. By a process of mutation, single strand break, mispairing, repair and replication, the 5bp sequence AGACA is duplicated on one allele. This process of slipped strand mispairing is believed to cause both duplication and deletion at points of repetitive sequence (Levinson & Gutman, 1987).

6.2.2 Misalignment of repetitive DNA during homologous recombination

An alternative means of sequence duplication and deletion involves misalignment during homologous recombination. This may be the cause of the 28 bp deletion found in *CYP11B1* exon 2 of subject 6, which occurs in a 36 bp region between codons 104 and 114. In this region the sequence reads:

GAC AGC CTG CAT CCC CAC AGG ATG AGC CTG GAG CCC

In the above sequence the repetitive bases are underlined. There are 3 AGCC motifs present within the 36 bp region, but the sequence AGC CTG CAT CCC at the 5' end of this stretch of DNA only differs from the 3' end, AGC CTG GAG CCC, by the two underlined bases. This could lead to misalignment during homologous recombination between two *CYP11B1* genes. Such a misalignment would, however, only have deleted 21 bp. The removal of the last seven base pairs could have been achieved by slipped strand mispairing of the third AGCC repeat, this time without strand breakage to generated a single strand loop of seven bases, AGC CTG G, which was then excised. On repair and replication an allele with a 28 bp deletion would have been generated. An alternative explanation, which is perhaps more likely, is that there was a straight misalignment at recombination between the first AGCC and the third AGCC motif.

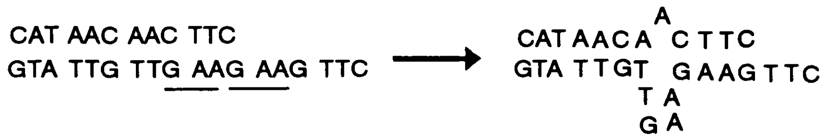
Once again *CYP11B2* has a different sequence in this region, removing the middle AGCC repeat and so would not be as prone to this mutation as *CYP11B1*. There is a possibility that increased repetitive sequence in *CYP11B1*, relative to *CYP11B2*, may be an underlying cause of greater mutation within certain regions of the *CYP11B1* gene.

The presence of repetitive DNA is presumed to be a major cause of duplication and deletion of DNA sequences and is a process affecting the *CYP11B* locus.

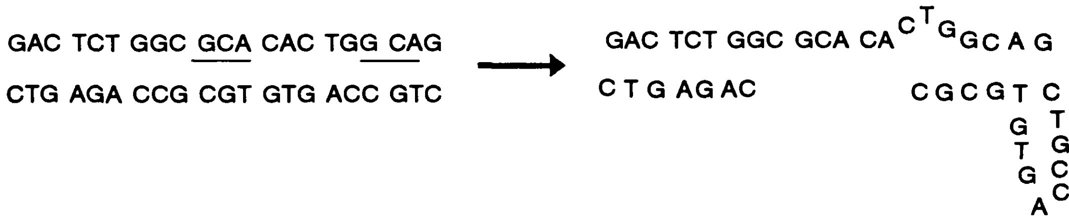
Small duplication and deletion events are also prominent among the deleterious mutations discovered in *CYP17*. Three duplications and two deletions have been reported (Yanase *et al* 1991; Imai *et al*, 1993), representing a third of the reported mutations in *CYP17*. Repetitive sequences also occur at the points of most of these mutations as shown in Figure 6.2b.

Figure 6.2b: Repetitive sequences near points of deletion and duplication reported in *CYP17*, and possible causes of mutations.

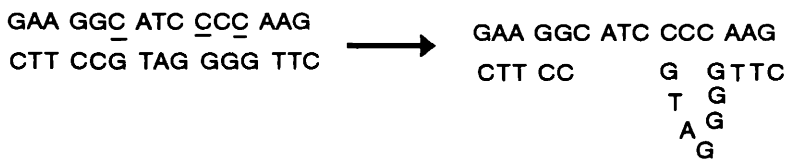
1. 3 bp of codon 54 due to mispairing during replication or repair



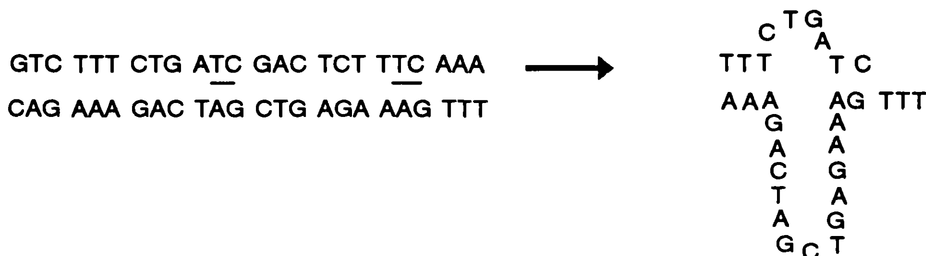
2. 7bp duplication at codon 120 due to single strand break, mispairing and repair



3. 4bp duplication at codon 480 due to single strand break, mispairing and repair



4. 9bp deletion of codons 487-489 due to mispairing during replication or repair



In the above diagrams underlined sequences represent repetitive bases that could have caused mispairing that lead to the deletions and duplications.

6.2.3 Deamination of methylated CpG dinucleotides in the *CYP11B* genes

Cytosine residues of CpG dinucleotides are often methylated in eukaryote genomes (Bird, 1980). Virtually all methylation in higher eukaryotes occurs at the 5-carbon position of cytosine, whilst prokaryotes and some lower eukaryotes contain both 5-methylcytosine and 6-methyladenine. Around 90% of DNA methylation in animal cells occurs at CpG dinucleotides, 10% involving methylation of other C residues.

In a simple chemical reaction 5-methylcytosine can be deaminated and converted to thymine (Ehrlich & Wang, 1981). For this reason methylated CpG dinucleotides (^{m}CpG) are hotspots for mutations involving C→T transitions, or G→A transitions, which are the result of deamination of methyl cytosine on the complementary strand. However, all C→T or G→A transitions do not occur by this process, as such mutations can also occur by simple misincorporation.

The number of CpG dinucleotides present in a gene and the degree to which they are methylated is believed to affect the rate at which transition mutations arise in a gene (Bird, 1980; Ehrlich & Wang, 1981). The relatively high level of CpG dinucleotides in *CYP11B1* (3% as opposed to 1.5% in the general human genome) has been cited as a reason for mutation at this locus (Curnow *et al*, 1993). These figures consider the whole gene. In the exonic sequence of *CYP11B1* there are 56 CpG dinucleotides, which means 56 potential points for mutation if all can be methylated.

In the *CYP11B2* exonic sequence there are 57 CpG dinucleotides of which 42 are shared with *CYP11B1*. Comparison of the two gene sequences where only one gene has a CpG dinucleotide revealed that 12 of the *CYP11B1* specific CpG dinucleotides correspond to points where *CYP11B2* sequence could have been derived from *CYP11B1* sequence by deamination of a CpG. Similarly, 12 out of the 15 *CYP11B2* specific CpG dinucleotides correspond to points where *CYP11B1* sequence could have been derived from similarly mutated *CYP11B2* sequence. This may suggest that the ancestral gene of *CYP11B1* and *CYP11B2* was actually a mixture of the two sequences and that deamination of ^{m}CpG dinucleotides has played a large part in the divergence of the *CYP11B* gene sequences.

The two linked mutations reported to cause CMO II deficiency, R181W and V386A (Pascoe *et al*, 1992b) could both have been caused by deamination of a ^{m}CpG dinucleotide. Four of the eleven reported mutations in CYP11B1, T318M,R374Q, R374Q and R448H are also of this type. Three mutations located in this study: ACG→ACA, T318T, CGC→CAC, R427H and CGC→CAC, R470H are all possibly due to this phenomenon.

6.2.4 Deamination of methylated CpG dinucleotides in *CYP17*

The pattern of CpG dinculeotides in *CYP17* and *CYP21* is also of interest. There are 59 exonic CpG dinculeotides in *CYP21*, a similar number to that in the *CYP11B* genes, but only 38 in *CYP17* (Appendix 7). These two genes are thought to have descended from the same ancestral gene (Picado-Leonard & Miller, 1987), in a similar fashion to the two *CYP11B* genes, yet have diverged to a such an extent that there is only a limited sequence homology.

The fact that *CYP17* contains less exonic CpG dinucleotides may be a contributing factor to the reduced level of mutation at this locus relative to both *CYP11B* genes and *CYP21*. A bigger factor in reducing transition mutations at the *CYP17* locus is that the gene is actively expressed in gonadal tissue: actively transcribed genes are believed to be a poor substrate for DNA methylation enzymes and are not fully methylated, possibly as a result of transcription factors and stretches of single stranded DNA interfering with methylase binding (Ehrlich & Wang, 1981). This being so, *CYP17* would be expected to be undermethylated in comparison with the *CYP11B* genes and *CYP21*, which are not expressed at high level in the gonads. As the gonadal tissue gives rise to the zygotes, the reduction in transition mutations in *CYP17* would be expected to translate into less inherited cases of 17α-hydroxylase deficiency compared to genes not actively transcribed in the gonads.

Three of the reported deleterious mutations in *CYP17* are possibly due to deamination of a CpG dinucleotide R239X, R440H and R496C (Ahlgren *et al*, 1992; Fardella *et al*, 1994a; Yanase *et al*, 1991), but these mutations could also be the result of misincorporation of nucleotides.

6.2.5 Misincorporation of nucleotides during repair or replication

Many deleterious mutations are caused by single base misincorporations arising from faulty repair or replication. Discounting mutations that may have arisen by deamination of ^{m}CpG dinucleotides, in the study presented here six mutations were due to misincorporation of nucleotides: G267R, G267D, Q356X, V458A, C494F in CYP11B1 and E194X in *CYP17*. Five out of the eleven deleterious mutations reported in CYP11B1 fall into this category (W116X, K174X, Q328X, Q356X and V441X). Of the sixteen reported *CYP17* mutations six are due to nucleotide misincorporation (W17X, Y64S, S106P, E194X, Q461X, and P468T).

The fact that within exons *CYP11B1* and *CYP11B2* differ only by nucleotide substitutions, allows the relative rates of misincorporation of nucleotides in the two genes to be estimated, assuming that sequence differences have arisen by divergence from a common ancestral gene. Of the eighty or so points where the exonic sequences of *CYP11B1* and *CYP11B2* differ, 62 are possibly due to misincorporation of nucleotides during repair or replication of DNA. Over a third of these (34%) involve C→T or T→C transitions, which could be due to simple misincorporation, but possibly also deamination of methylcytosine residues. Transitions involving purines are rarer, (at 19%). Transversions from G→C or C→G account for 24% of apparent misincorporations, whilst the other transversions are much lower occurring at 10%, 8%, and 5% for G→T/ T→G, T→A/ A→T and C→A/ A→C substitutions respectively. The differences between *CYP11B1* and *CYP11B2* are shown in Appendix 6. The rate of purine transitions may be somewhat higher if the CpG deaminations excluded here were partly due to simple misincorporations. From the above figures it would appear that certain misincorporations are much more common than others and that GC rich regions of DNA may be more prone to misincorporation.

The efficiency of repair may also vary depending on the different mismatches present, which could affect the occurrence of certain misincorporations (Bessman & Reha-Krantz, 1977; Petruska & Goodman, 1985). The rates of misincorporation are currently being extensively analyzed and are discussed in detail in Cooper & Krawczak (1993).

In the analysis carried out in this study, it was noted that similar types of nucleotide change cluster in small regions. In codons 67, 68 and 74, for instance the *CYP11B1* and *CYP11B2* genes differ by four presumed misincorporations involving G→A/ A→G transitions and only eight more such changes are found in the rest of the coding region. It is possible that patches of repair are carried out by enzymes with faulty proofreading capability, so that small regions develop misincorporations of a particular kind. Although the events causing mutations are random the distribution of different types of mutation is not and may be highly dependent on the sequence at the point of change.

6.2.6 Multiple mutations in *CYP11B1*

Multiple mutations of the *CYP11B1* gene were found in two patients, subjects 1 and 6. Most reported deleterious mutations are rare and isolated events. In subjects 1 and 6, more than one potentially harmful mutation had occurred on the same allele. In both cases a misalignment event, leading to small deletion or insertion of DNA, was found in combination with a single base substitution leading to an amino acid change. This implies that two separate mutation events had affected the same allele. In order for this to happen a mutant allele would have to be maintained in a population until a second mutation occurred. The question of whether the first mutation increases the risk of generating a second is a point worth considering. One hypothesis would be that a mutation within *CYP11B1*, giving rise to a non-functional protein, would prevent transcriptional control mediated by the enzyme product. In this hypothesis an unusually high rate of transcription would follow, with an increased chance of DNA strand breakage, or misalignments between repetitive sequences. This would only occur in individuals homozygous for the first defect, or in cells, such as precursors of gametes, that only contain single copies of chromosomes.

The high CpG content and perhaps a high degree of methylation of *CYP11B1* may be in part responsible for increasing the chance of mutations occurring. The number of repetitive elements may be higher in *CYP11B1* than in some other genes. However, it is also possible that the amino acid changes found do not have a deleterious effect. No other group has reported multiple deleterious mutations on a single allele of the *CYP11B1* gene.

6.2.7 Gene conversion and non-homologous recombination

The two *CYP11B* genes arose from the same ancestral sequence and consequently the apparent gene conversions from *CYP11B1* to *CYP11B2* may in fact be points at which the ancestral duplicated sequence has not mutated to the *CYP11B1* type in certain populations. This could well be the case for the two gene conversions that constitute common variants in introns 3 and 4. *CYP11B1* specific sequence, as found in Type A individuals (see section 4.6.2) and previously published sequences, would represent mutation away from the ancestral gene that gave rise to both *CYP11B* genes. Conversely, Type B individuals have a *CYP11B1* gene that has not diverged from the ancestral sequence at certain points that appear as gene conversions to *CYP11B2*.

For the intron 3 gene conversion, involving only 8 bp, there are *CYP11B1* sequences to either side. The explanation that mutations gradually arose in a stretch of *CYP11B2* like sequence may be more plausible than an 8 bp patch of gene conversion. On the other hand, the mutation of the intron 3 area to a give the *CYP11B1* pattern involves five base differences and could have been the result of a single misalignment event or error in repair. In this case the intron 3 polymorphism could well be a small gene conversion, inserting *CYP11B2* sequence back into a mutated region of DNA.

The smaller the gene conversion is, the less plausible they become. A single base change corresponding to the related gene could well be just a chance mutation. The fact that both genes are mutating away from the ancestral predecessor, and that there is a possibility of interchange of sequence between the two, confuses the picture of what has occurred. This being so gene conversion may not be as common as it would at first appear. Evidence that it has occurred is supported by the fact that subjects 13 and 14 possess a clear *CYP11B1* type sequence in intron 4 of their *CYP11B2* and examples of *CYP11B1* to *CYP11B2* gene conversion that are not found in other individuals studied. That they have unique gene conversions, whilst belonging to the general polymorphic group B suggests that gene conversion is an ongoing process in the evolution of duplicated genes.

The evidence for gene conversion in 21-hydroxylase deficiency is much clearer. Here the changes that occur insert sequences from a pseudogene into exonic regions of *CYP21* and abolish enzyme function. It would seem logical that the same process of sequence exchange occurs between the *CYP11B* genes, but possibly not as often (as the genes are further apart: 45 kb as opposed to 30 kb in the case of the *CYP21* locus) and with less effect, because both genes produce functional enzymes that are 93% identical in amino acid sequence. The presence of repetitive sequences involved in promoting recombination, with homology to bacterial chi sequence (GCTGTGGG), has also been cited as a possible cause of specific gene conversion at the CYP21 locus. Such sequences do not appear to be present in the *CYP11B* genes. Gene conversion, of exons 3 and 4 of *CYP11B2* to *CYP11B1*, has recently been cited as causing a case of CMO II deficiency (Fardella *et al*, 1994b), but no other cases have been reported to date.

Non-homologous recombination was not found in our sample of subjects. This process, generating a *CYP11B1/CYP11B2* chimaeric gene has been shown to cause glucocorticoid remediable hyperaldosteronism (Pascoe *et al*, 1992a; Lifton *et al*, 1992). Initially subjects 15 and 16, who exhibited an autosomal dominant form of adrenal-related hypertension, were suspected to have GRH, but the typical form of GRH was excluded by Southern blot analysis. Another explanation is currently being sought to explain the condition of these two subjects, including the possibility that gene conversion is involved.

6.3 Mutation rate in the *CYP11B*, *CYP17* and *CYP21* genes

The most commom form of CAH is due to 21-hydroxylase deficiency, which affects around 1 in 12,500 live births (New *et al*, 1989). This contrasts with 1 in 100,000 for 11ß-hydroxylase deficiency (White & Pascoe, 1992). CMO I and CMO II deficiency, involving loss of aldosterone synthase function are very rare, as are cases of 17α-hydroxylase deficiency, where only just over a hundred cases have been described in the literature. A number of factors must underlie these marked differences in apparent mutation rates of the three cytochrome P450 genes involved in these syndromes.

21-hydroxylase is situated in tandem with a pseudogene, only 30 kb apart, on chromosome 6, within the HLA major histocompatibility complex (Higashi *et al*, 1986). This unique location has had a considerable effect on the generation of mutations leading to 21-hydroxylase deficiency. Approximately 25% of mutations causing 21-hydroxylase deficiency involve complete loss of the functional gene due to unequal crossover (Rumsby *et al*, 1986; Werkmeister *et al*, 1986). This arises due to misalignment, between highly homologous repeated C4 genes 5' of both pseudogene and functional *CYP21*.

Most of the remaining mutations discovered in CYP21 are due to gene conversions altering the functional to the non-functional sequence of the pseudogene (Harada *et al*, 1987; Amor *et al*, 1988; Globerman *et al*, 1988; Higashi *et al* 1988; Chiou *et al* 1990). The bulk of mutations within *CYP21* are thus due to the unusual structure of the *CYP21* locus in which a pseudogene is situated in tandem with a functional gene and both are flanked at the 5' end with a second set of duplicated genes. The rate of spontaneous mutations in *CYP21* is, then, probably of a similar order to that in *CYP11B1*.

The location of *CYP21* in the MHC region, which is one of the most repetitive and changeable sections of the genome, would not be envisaged as a particularly good site for a gene involved in development. The retention of the pseudogene is also curious. Evolutionary theory would suggest an advantage to this arrangement. One possibility

is that there is a heterozygote advantage gained from 21-hydroxylase deficiency. Hirsutism, short stature, excess weight gain and possibly aggression are all factors that could be of benefit in adaptation to a cold environment or one in which a feast or famine existence prevails. Cortisol increases available blood sugar by stimulating gluconeogenesis from amino acids. Lack of cortisol reduces this effect, effectively slowing metabolism down. Having individuals in a population who carried the potential for reduced cortisol production, however mildly, may have just tipped the balance between survival and extinction and been advantageous to a primitive ancestral group. The same heterozygous advantage may also apply to 11ß-hydroxylase deficiency.

The existence of peripheral 21-hydroxylase enzymes could well mean that total CYP21 deficiency does not lead to total loss of cortisol production (Mellon & Miller, 1989; Speiser *et al*, 1991). A separate P450 gene, CYP2C5, encodes a liver steroid 21-hydroxylase enzyme in rabbits (Pendurthi *et al*, 1990) and similar enzymes may be present in man.

The location of a gene for a extracellular matrix protein on the opposite strand of the pseudogene (Morel *et al*, 1989), overlapping with its 3' end, may also provide a reason for the retention of a potentially harmful stretch of DNA in the gene pool. The pseudogene may be maintained so as not to lose this other more essential protein.

Gene conversion and non-homologous recombination at the *CYP11B* locus on 8q do not have the same results as at the *CYP21* locus. Unequal crossover can generate the chimaeric, functional gene that gives rise to glucocorticoid remediable hyperaldosteronism, but this is rare when compared to *CYP21* rearrangements. The fact that CYP11B1 and CYP11B2 are both functional proteins means that gene conversion events between the two may not abolish function, as is more often the case of gene conversions between a non-functional pseudogene and a functional homologue. Secondly, the deletion of *CYP11B* genes has not been reported. Differences between the 5' regions of *CYP11B1* and *CYP11B2* may help to reduce the chance of deletion by unequal crossover and the two genes are 45 kb apart as opposed to 30 kb in the case of *CYP21* and *CYP21P*. Complete loss of *CYP11B2* may be

lethal, due to loss of aldosterone and 18-hydroxylated steroids. In CMO II deficiency 18-hydroxylated products from mutant genes may still provide some mineralocorticoid activity. Deletion of *CYP11B1*, however, might be expected to occur and be maintained in heterozygotes, but has not been reported.

The fact that introns in the *CYP11B* genes are much longer than in CYP21 is also of interest. With short introns any misalignment during recombination, or gene conversions that occur are more likely to involve exons. The introns of *CYP21* range from 83-282 bp in length. In contrast the smallest introns of *CYP11B* genes are 81 and 135 bp (introns 8 and 3), comparable to *CYP21* introns, but the rest vary from 284-806 bp in length. In the study presented here all gene conversions appeared in introns and consequently are unlikely to lead to loss of enzyme function. This is in contrast to 21-hydroxylase deficiency, where gene conversions commonly affect the exons of the functional gene. The notable exception to this is the intron 2 gene conversion mutation reported in *CYP21* (Higashi *et al*, 1988), which affects splicing and is a common cause of 21-hydroxylase deficiency (Rumsby *et al*, 1992b). However, this mutation occurs within the largest intron of *CYP21* (282 bp in length).

The low level of mutation within *CYP17*, relative to *CYP21* and *CYP11B1* genes, is affected by some of the factors mentioned above. *CYP17* is a single copy gene. This reduces the opportunity for non-homologous recombination and gene conversion. Mutations occurring in *CYP17* are of the spontaneous kind, only influenced by repetitive sequences within the gene, the degree of deamination of methylated CpG dinucleotides and random errors in repair, replication or homologous recombination. The fact that *CYP17* is actively transcribed in gonads, whereas *CYP21* and *CYP11B* genes are not, may also have an effect on the types of mutation present in *CYP17.* Actively transcribed genes would be expected to be less prone to methylation and therefore mutation by deamination of methylcytosine residues (Bird, 1980; Ehrlich & Wang, 1981). Actively transcribed genes are more prone to breakage during the transcription process, but are believed to be repaired and proof read more efficiently than non-transcribed genes (Bohr *et al*, 1985; Vos & Waulthier, 1991), which could account for a low level of inherited mutations in *CYP17.* Selective pressures against mutation in *CYP17* may also be stronger than for the *CYP21* or *CYP11B1* genes.

The incidence of deleterious mutations within *CYP11B2* is again very low. In part this may be due to salt loss being lethal in neonates. The same applies to 21-hydroxylase deficiency, but here, as discussed above, aldosterone synthesis may not be completely abolished, even in complete absence of CYP21, because of extra-adrenal 21-hydroxylase enzymes (Speiser *et al*, 1991). There are probably strong selective pressures against retaining *CYP11B2* mutations in a population.

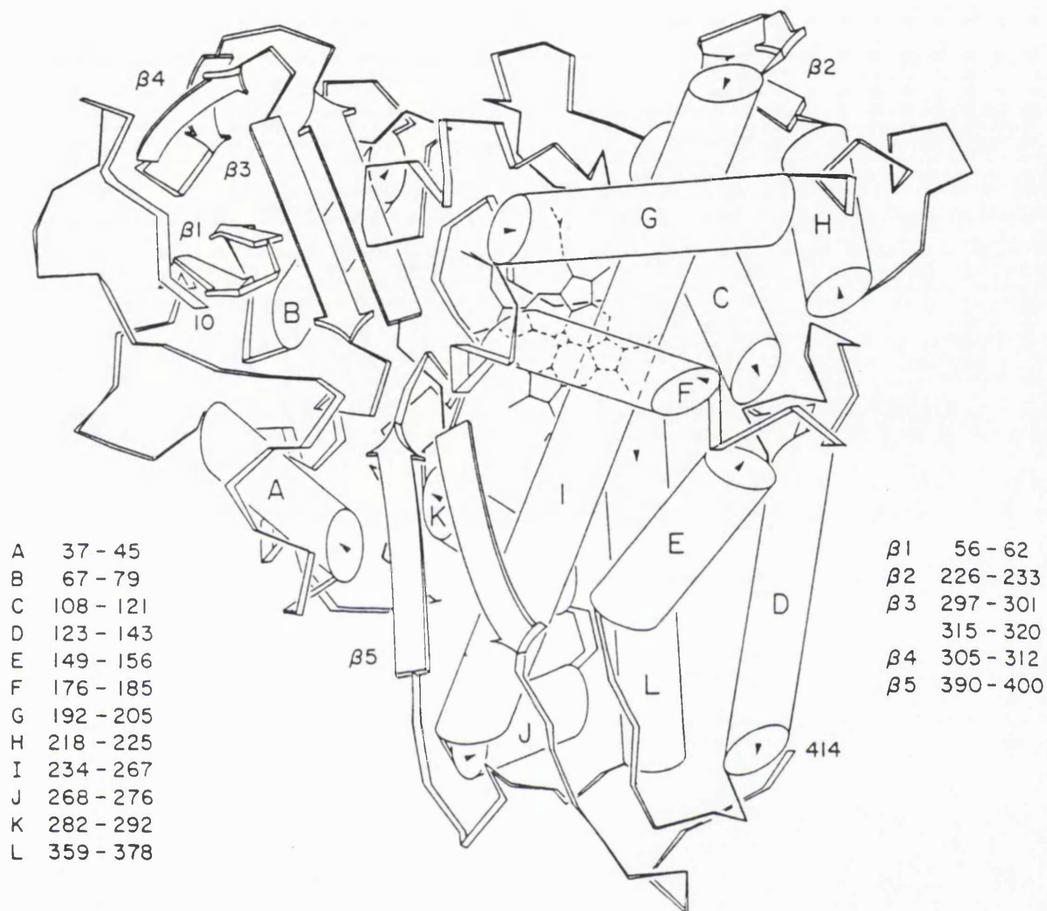
The difference between mutation rates of *CYP11B1* and *CYP11B2* may well be mirrored by their functions. CYP11B1 has a single function of adding an 11ß-hydroxyl group, whereas CYP11B2 carries out 11ß-hydroxylation, 18-hydroxylation and the formation of an aldehyde group at C-18. There may be more chance for amino acid change in CYP11B1 without totally destroying function, whilst CYP11B2 may be very sensitive to change. The mutation of *CYP11B1* may be tolerated to a greater extent than mutation of *CYP11B2*. There may well be subtle differences in the degree of repetitive sequence between the *CYP11B* genes, as discussed in section 6.2.1 and 6.2.2, and other sequence differences that lead to less mutation at the *CYP11B2* locus.

6.4 Amino acid changes in CYP11B1, CYP11B2 and CYP17

Insufficient time was available to express mutations in tissue culture, but alignment of amino acid sequences with those of the two crystallized bacterial P450 enzymes was carried out and proved useful in hypothesizing the effect of certain amino acid changes. The proposed structure of CYP101 is shown in Figure 6.4a and is followed by the alignment of amino acid sequences from CYP101, CYP11A, CYP11B1, CYP11B2, CYP17, CYP21 and CYP102 (Figure 6.4b).

From this alignment it would appear that the amino acid changes characterized in this study may well affect substrate binding and the P450 structure rather than abolish haem function. In the following section the proposed positions of mutations are given, related to the standard model shown in Figure 6.4a.

Figure 6.4a : Proposed three dimensional structure of CYP101, a bacterial P450 from *Pseudomonas putida*, determined by X-ray crystallography (taken from Poulos *et al*, 1985)



Above is shown the proposed structure of CYP101. Cylindrical sections labelled A-L represent α-helices, whilst paired arrows represent anitparallel ß-pairs. The amino acids forming these structures are listed above.

Figure 6.4b : Alignment of P450 amino acid sequences; CYP101, CYP11A, CYP11B1, CYP11B2, CYP17 CYP21, and CYP102.

```
(a)------------------------------------------------TTETIQSNANLAPLPPHVPEHLVFDFDMYNPSNLSA---------- 36
(b)MLAKGLPPRSVLVKGYQTFLSAPREGLGRLRVPTGEGAGISTRSPR-------------------PFNEIPSPGDNGWLNLYHFWRETGTHKVH 75
(c)MALRAKAEVCMAVPWLSLQRAQALGTRAARVPRTVL-----------------------------PFEAMPRRPGNRWLRLLQIWREQGYEDLH 65
(d)MALRAKAEVCVAAPWLCLQRARALGTRAARAPRTVL-----------------------------PFEAMPQHPGNRWLRLLQMWREQGYEHLH 65
(e)MWELVALLLLTLAYLFWPKRRCPGAK------------------------------------------YPKSLLSLPLVGSLPFLPRHGHMH 50
(f)MLLLGLLLLPLLAGARLLWNWWKLR------------------------------------------------SLHLPPLAPGFLHLLQPDLP 45
(g)-------------------------------------------------------------TIKEMPQPKTFGELKNLPLLNTDK----- 24

      A                 ß1           B                                          * C
(a)GVQEAWAVLQ-ESNVP-DLVWTRCNGGHWIAT-RGQLIREAYED-YRHFSSECPFI--PREAGEAY-DFIPTSMDP-PEQRQFRALANQVVGMP- 122
(b)LHHVQNFQKYG--PIYREKLG--NVESVYVIDPEDVALLFKSEGP--NPERFLIPPWVAYHQYYQRPI-GVLLKK-SAAWKKDRVALNQEVMAP- 161
(c)LEVHQTFQELG--PIFRYDLG--GAGMVCVMLPEDVEKLQQVDSL--HPHRMSLEPWVAYRQHRGHKC-GVFLLN-GPEWRFNRLRLNPEVLSP- 151
(d)LEMHQTFQELG--PIFRYNLG--GPRMVCVMLPEDVEKLQQVDSL--HPCRMILEPWVAIRQHRGHKC-GVFLLN-GPEWRFNRLRLNPDVLSP- 151
(e)NNFFKLQKKYG--PIYSVRMG--TKTTVIVGHHQLAKEVLIKKGKDF-SGRPQMATLDIASNNRK---GIAFAD-SGAHWQLHRRLAMATFALFK 136
(f)IYLLGLTQKFG--PIYRLHLG--LQDVVVLNSKRTIEEAMVKKWADF-AGRPEPLTYKLVSKNYPDLSL-GDYSLL---WKAHKKLTRSALLLG- 130
(g)PVQALMKIADELG-EIFKFEAPG-RVTRYLS-SQRLIKEAC-DES-RFDKN-L-SQALKFVRDFAGDGLFTSWTH-EKNWKKAHNILLP----- 105
      A           ß1      ß2      B         ß3          B'                  C

                D                                              E
(a)-----VVDKLENRIQELACSLIESLRPQ------------------------GQCNFTEDYAEPFPIRIFMLLAGL--------------- 169
(b)--EATKNFLPLLDAVSRDFVSVLHRRIKKAGSG--NYSGDISDDLFRFAFESITNVIF--GERQGMLEEVVNPEAQRFIDAIYQ---------- 239
(c)--NAVQRFLPMVDAVARDFSQALKKKVLQNARGS--LTLDVQPSIFHYTIEASNLALF--GERLGLVGHSPSASSLNFLHALEV---------- 229
(d)--NAVQRFLPMVDAVARDFSQALRKKVLQNARGS--LTLDVQPSIFHYTIEASNLALF--GERLGLVGHSPSASSLNFLHALEV---------- 229
(e)--DGDQKLEKIICQEISTLCDMLATHNG-QSIDISFPVFVAVTNVISLICFNTSYKN---GDPELNVIQNYNEG-------------------- 204
(f)--IRDSMEPVVEQLTQEFCERMRAQPG------TPVAIEEEFSLLTCSIICYLTF----GDKIKDDNLMPA--------------------- 189
(g)SFQQAMK--GYHAMMVDIAVQLVQKWER--------LNADE----HIEV-----------PEDMTRLTLDTIGLCGF----------------- 157
                 D                        ß4                  E

                             F                           G
(a)-PEE-----------------DIPHLKYLTDQMTR-PDGSM------TFAEAKEALYDYLIPIIEQRRQK-----------PGT--------- 217
(b)--------------MFHTSVPMLNLPPDLFRLFRTKTW-KDHVA---AWDVIFSKADIYTQNFYWELRQKGS--------VHHDYRGM------ 301
(c)--------------MFKSTVQLMFMPRSLSRWTSPKVW-KEHFE---AWDCIFQYGDNCIQKIYQELAFSRP--------QQYTSIVA------ 291
(d)--------------MFKSTVQLMFMPRSLSRWISPKVW-KEHFE---AWDCIFQYGDNCIQKIYQELAFTRP--------QHYTGIVA------ 291
(e)---------IIDNLSKDSLVDLVPWLKIFPNKTL------EKLKS---HVKIRNDLLNKILENYKEKFRS-----------DSITN--------- 261
(f)YYKCIQEVLKTWSHWSIQIVDVIPFLRFFPNPGL------RRLKQ---AIEKRDHIVEMQLRQHKESLVAG----------QWRD--------- 255
(g)------NYRFNSFYRDQPH-PFITSMVRALDEAMNKLQ--RANPDDP-AYDENKRQFQEDIKVMNDLVDKIIADRKAS-GEQSDD--------- 231
                               F                            G

    H              ß2                     I **                   J                     K
(a)DAISIVAN-------GQVNGPRI---TSDEAKRMCGLLLVGGLDTVVNFLSFSMEFLAKSP-EHRQELIER-------PE-------------R 280
(b)LYRLLGDSK---------------MSFEDIKANVTEMLAGGVDTTSMTLQWHLYEMARNL-KVQDMLRAEVLAARHQAQ---GDMATMLQLVPL 376
(c)ELLLNA-----------------ELSPDAIKANSMELTAGSVDTTVFPLLMTLFELARNP-NVQQALRQESLAAAASIS---EHPQKATTELPL 364
(d)ELLLNA-----------------ELSLEAIKANSMELTAGSVDTTAFPLLMTLFELARNP-DVQQILRQESLAAAASIS---EHPQKATTELPL 364
(e)MLDTLMQAKMNSDNGNAGP-DQDSELLDNHILTTIGDIFGAGVETTTSVVKWTLAFLLHNP-QVKKKLYEEIDQNVGFSRTPTIS---DRNRLLL 351
(f)MMDYMLQGVAQPSMEEGSG-----QLLEGHVHMAAVDLLIGGTETTANTLSWAVVFLLHHP-EIQQRLQEELDHELGPGASSSRVPYKDRARLPL 344
(g)LLTHML--NGKDPETGEPL-------DDENIRYQIITFLIAGHETTSGLLSFALYFLVKNP-HVLQKAAEEAARVLVDPVP----SYKQVKQLKY 312
    H                                   I                          J                  K

       K*  *       ß3* *       ß4        ß3                           *
(a)IPAACEELLRRF---SLVADGRIL--TSDYEFHGVQL-KK-GDQILLPQM--------LSGLDERENACPMHVDFSRQKVS------------- 346
(b)LKASIKETLRLHP--ISVTLQRYL--VNDLVLRDYMI--PAKTLVQVAI--------YALGREPTFFFDPENFDPTRWLSKDKNITYFR------ 451
(c)LRAALKETLRLYP--VGLFLERVA--SSDLVLQNYHI--PAGTLVRVFL--------YSLGRNPALFPRPERYNPQRWLDIKGSGRNFY------ 439
(d)LRAALKETLRLYP--VGLFLERVV--SSDLVLQNYHI--PAGTLVQVFL--------YSLGRNAALFPRPERYNPQRWLDIKGSGRNFH------ 439
(e)LEATIREVLRLRP-VAPMLIPHKA--NVDSSIGEFAVDK-GTEVIINL---------WALHHNEKEWHQPDQFMPERFLNPAGTQLISPSVS--- 431
(f)LNATIAEVLRLRP-VVPLALPHRT--TRPSSISGYDIP--EGTVIIPNL--------QGAHLDETVWERPHEFWPDRFLEPGKNSR--------- 417
(g)VFMVLNEALRLWP-TAPAFSLYA--K-EDTVL--GG--EYPLE-K-GDELMVL-IPQLH-RDKTIWGDDVEEFRPERFENPSAIPQHA------- 389
      K           ß5      ß6        ß7     ß8    K'
      *   *    * ***   *   *                           ß5
(a)HTT--FGHGS--HLCLGQHLARREIIVTLKEWLTR--IPDFSIAPG-AQIQHKSGIVSGV-----QALPLVWDPATTVRV-------------- 414
(b)NLG--FGWGV--RQCLGRRIAELEMTIFLINMLEN--FRVEIQHL--SDVGTT--F--NLILM--PEKPISFTFWPFNQEATQQ----------- 521
(c)HVP--FGFGM--RQCLGRRLAEVEMLLLLHHVLKH--LQVETLTQ--EDIKMV--Y--SFILR--PSMCPLLTFRAIN----------------- 503
(d)HVP--FGFGM--RQCLGRRLAEVEMLLLLHHVLKR--FLVETLTQ--EDIKMV--Y--SFILR--PGTSPLLTFRAIN----------------- 503
(e)YLP--FGAGP--RSCIGEILARQELFLIMAWLLQR--FDLEVPD-DGQLPSLEG-I-PKVVFLID--SFKVKIK--VRQAWREAQAEGST----- 508
(f)ALA--FGCGA--RVCLGEPLARLELFVVLTRLLQA--FTLLPSG---ALPSLQP-L-PHCSVIL--KMQPFQVR--LQPRGMGAHSPGQNQ---- 494
(g)FKP--FGNGQ--RACIGQQFALHEATLVLGMMLKH--FDFE-DHTNYELDIKET-L-TLKPE--G-FVVKAKSKK-IPLGGIPSPSTEQSAKKVR 471
                      L              ß9          ß10    ß11     ß12
```

Seven P450 enzyme amino acid sequences are aligned above; (a) CYP101, from *Pseudomonas putida*, (b) human CYP11A, (c) human CYP11B1, (d) human CYP11B2, (e) human CYP17, (f) human CYP21 and (g) CYP102 from *Bacillus megaterium*. Numbers at the end of lines give the residue number ending that line. For CYP101 and CYP102, regions underlined are those which are believed to form secondary structures: regions lettered A-L form helices, those lettered ß1-ß5 form ß pairs (alignment for CYP101, CYP17 and CYP21 taken from Nelson & Strobel (1987), CYP101 secondary structure data taken from Poulos *et al* (1985), CYP102 secondary structure data taken from Ravichandran *et al* (1993)). Asterisks denote highly conserved amino acids (see Appendix 8 for details).

The G267R mutation in CYP11B1 is predicted to occur at the N-terminal of helix G, which lies above the substrate binding pocket. An arginine at this position could extend its side chain into hydrophobic regions of the protein involved in substrate binding and disrupt function. It could also form charge interactions with acidic amino acids and disturb the conformation of the protein, making it susceptible to proteolysis. The G267D mutation may have the same type of effect in altering substrate binding ability or changing structure by charge interaction with basic amino acids.

The arginine at position 427 is highly conserved in steroidogenic P450 enzymes and is located in a loop region immediately above the haem group. Substitution with histidine may disrupt a charge interaction required to maintain enzyme structure at this point.

The V458A mutation may represent a polymorphism. A valine was initially reported at this position in normal *CYP11B1* (Mornet *et al*, 1989), but the sequence entry in the Human Genome Project databank (Accession number M32878 J05140), which is taken from the same reference, has an alanine codon at this point. The change is relatively conservative, however, and may have little or no effect on enzyme activity.

The C494F mutation, found in subjects 10, 13 and 14, occurs 10 amino acids from the C-terminal of CYP11B1. The two bacterial P450 structures that have been determined by X-ray crystallography have a ß-pair loop near this point, which is in contact with the F-G loop and forms important contacts for the substrate and redox partner binding sites. A change from cysteine to an aromatic phenylalanine could well alter the positioning of the ß-loop equivalent in CYP11B1 and affect substrate or adrenodoxin binding.

The R470H mutation present in the CYP11B2 of subjects 13 and 14 is positioned at the opposite end of helix L to the haem. Several P450 enzymes, including CYP11B1, have histidine at this point, which suggests that it would not have a radical effect on CYP11B2 function. In cases of CMO II deficiency 18-hydroxylated products can be produced, but the final step: converting the 18 and 11-hydroxyl groups to a hemiacetal and then aldehyde group, cannot be achieved. This is also true for 11ß-

hydroxylase, which can produce low levels of 18-hydroxy compounds, but not aldosterone. There is a possibility, however, that differences at the C-terminal end of helix L could alter the position of the haem subtly so that specific substrates are positioned correctly for the P450 reactions to occur. Verification that R470H has an effect on CYP11B2 production of aldosterone is required.

The excess 18-hydroxylated steroids produced by subject 16 lead to the question of whether key amino acids in CYP11B1 had become altered to CYP11B2 on a single allele, so that 18-hydroxylated products were produced more efficiently by the 11ß-hydroxylase enzyme. SSCP analysis showed exon 4 of *CYP11B1* to be unique for both subjects 15 and 16. The only amino acid difference between CYP11B1 and CYP11B2 enzymes in the region encoded by exon 4 occurs at amino acid position 248, which is threonine in CYP11B1, but isoleucine in CYP11B2. Comparison with CYP102 and CYP101 suggests that T248 in CYP11B1 is part of a flexible loop involved in substrate interaction. A mutation of a single amino acid from F209 to L209 at the equivalent point of this F-G loop was sufficient to alter substrate specificity from coumarin to testosterone in mouse microsomal P450coh (Lindberg & Negishi, 1989). Sequence analysis of these patients is incomplete, only one allele having been cloned, which was normal in exon 4.

The effect of the change I248T, in CYP11B2, has been determined by Fardella *et al* (1994b), as this was one of the amino acid substitutions brought about by the gene conversion event in a patient with CMO II deficiency. In expression studies this mutation was shown to markedly reduce 18-hydroxylation activity of CYP11B2. It could therefore be hypothesized that the reciprocal change, T248I, in the CYP11B1 enzyme, may well increase the 18-hydroxylation activity of the CYP11B1 enzyme, which would still be highly transcribed in response to ACTH.

The five amino acid changes reported by other groups to cause loss of CYP11B1 enzyme function: T318M, R374Q, R384Q, V441G (Curnow *et al*, 1993) and R448H (White *et al*, 1991), all involve loss of haem function or are situated close to the haem. The T318M mutation alters an amino acid involved in passing protons to the haem group, which is necessary for catalysis of the P450 mediated splitting of

molecular oxygen. R384 of CYP11B1 is involved in binding of a haem propionate. The amino acid H373 of CYP17 corresponds to R384 of CYP11B1 and the mutation H373L has been reported to cause loss of CYP17 function (Monno *et al*, 1993). R448 in CYP11B1 also contacts the haem. The deleterious mutation R440H in CYP17 (Fardella *et al*, 1994a) corresponds to the R448H mutation in CYP11B1. Mutation of these amino acids presumably reduces or abolishes haem binding ability. R374 of CYP11B1 is also highly conserved and may be involved in maintaining enzyme structure close to the haem. The mutation V441G may affect the positioning of the loop above the haem.

The R181W mutation in CYP11B2, which causes CMO II deficiency, occurs in a region of the protein corresponding to the 3' end of helix D in the CYP102 model. The alteration of a positively charged residue to a hydrophobic aromatic residue presumably affects the positioning of helix D. The tryptophan may associate with other hydrophobic residues to shift the position of helix D and this could alter the relative positions of helix E and the vital F-G loop, which is believed to interact with the ß-pair region at the C-terminal and effect substrate specificity.

The remaining amino acid changes that have been reported in CYP17 (deletion of the phenylalanine at position 53 or 54, Y64S, S106P, duplication of I112, P468T, deletion of amino acids 487-489 and R496C) probably affect enzyme structure and substrate binding. The loss of a phenylanine, possibly in the region of helix A, could lead to altered conformation close to the substrate binding site, as could the change of tyrosine to serine at amino acid position 64. The serine at position 106 occurs between helix B and helix C, and the change to proline may alter conformation just before helix C, which carries a contact to the haem. The isoleucine at position 112 of CYP17 is believed to sequester the hydrophobic end of the substrate into the active site, suggesting that duplication at this point may affect substrate binding (Imai *et al*, 1993). The final three mutations all occur in the ß-pair region of the C-terminal, where changes could alter the position of the ß-loop required for substrate and redox partner binding.

Tables showing equivalent positions of key amino acids in the seven P450 enzymes aligned in this thesis are given in Appendix 8.

6.5 Evaluation of methods

The advent of PCR amplification of DNA has proved immensely useful to molecular biologists, in that specific stretches of DNA matching human genomic sequence can be obtained in a matter of hours. As such it amounts to a revolution in DNA investigation. The advent of sequencing automation, allowing mutiple samples to be prepared and analyzed, has also meant that the sequencing of an entire gene can be carried out in a matter of days as opposed to months. The ability to sequence directly from PCR product has made advances in sequence determination follow at a rapid pace. However, these developments all represent state of the art technologies. Manual sequencing of large stretches of DNA from multiple subjects remains a time consuming, and often frustrating process.

In light of difficulties with cloning and obtaining clear DNA sequence, SSCP proved to be very useful. Rather than sequencing large regions of DNA, particular regions of genes containing base differences were first highlighted by SSCP analysis. Size differences were easily identified by the high resolving power of PAGE and SSCP techniques. Some single base changes gave very subtle SSCP differences and only appeared after running gels of different compositions under varying electrophoresis conditions. Other single base changes produced very pronounced SSCP differences. The high degree of sequence variation in the sample of subjects included in this study produced many unique SSCP patterns. It should be stressed, however, that many of these changes represented intronic differences or silent mutations that had no effect on enzyme function. To restrict the number of non-exonic mutations, it is recommended that the region of intronic DNA included in PCR fragments for SSCP analysis should be limited to include only intronic sequences required for splicing.

Certain sequence changes were not detected by SSCP, including the C494F mutation of CYP11B1 and the R470H mutation of CYP11B2 in subjects 13 and 14. Similarly the intron 1 mutation was not responsible for SSCP differences seen in DNA from patients 6 and 10, as other sequence differences were also present. The intron 1 change represented a two base gene conversion, loss of five bases and duplication of four bases: a net loss of 1 bp and many changes that one would expect to give a

difference in SSCP pattern. These changes occurred near the 5' end of the PCR fragment, which implies that mutations at certain points on a DNA fragment have less effect on conformation. This was certainly the case with a gene conversion in intron 4, which was not clear on SSCP analysis of exon 4, where the mutation was at the 5' end of a 96 bp fragment, but was shown when SSCP analysis was carried out after digestion of a 784 bp fragment containing exon 4 and 5 with *Rsa*I enzyme. The digestion product of 330 bp had the gene conversion in the middle of the fragment. Carrying out different enzyme digests on PCR products and the use of additional primers to amplify overlapping regions could well improve the detection of single base changes. In addition the selection of suitable restriction enzymes also allows localization of sequence change to the 5' or 3' end of PCR fragments.

The development of non-isotopic SSCP removes the need for disposal of radioactive waste and makes the technique safer for the operator. Many groups have successfully used silver staining, including Chaubert *et al* (1993). Automated systems, such as the PhastSystem™ (Pharmacia, LKB Biotechnology, Uppsala, Sweden) are also currently in use. This system has the advantages of commercially available preformed gels, which can be rapidly electrophoresed in a controlled environment, coupled with an extremely sensitive silver stain that can visualize both single and double stranded DNA. It has been used successfully in the location of mutations within the human p53 gene involved in cancer (Mohabeer *et al*, 1991), in the study of Tay-Sachs disease (Ainsworth *et al*, 1991) and in the phenylalanine hydroxylase gene (Dockhorn-Dworniczak *et al*, 1991). This system probably represents the state of the art for SSCP analysis.

The disadvantages of PCR-SSCP analysis mainly lie in the production of differences due to PCR artifacts, which create false positives. This problem was not found in the study carried out, but has been experienced in other studies carried out in our laboratory. The inclusion of family members, to confirm the presence of SSCP differences and repeating SSCP from a separate initial PCR helped to eradicate such problems. Running samples from sixteen individuals allowed common polymorphisms to be detected. In further studies it would be desirable to run samples known to give differences, to check the resolution of gels used.

Optimising PCR for SSCP analysis was essential, so that only specific products were produced. The selective amplification of *CYP11B1* and *CYP11B2* DNA, followed by reamplification using a small amount of purified specific template DNA, allowed a good yield of specific DNA sequence for SSCP analysis in the study carried out. As an initial screening method SSCP proved very successful.

Obtaining the sequence of regions believed to carry base changes was not so simple. In the case of subject 14, six changes in *CYP11B1* all proved to be non-reproducible and presumably due to PCR or cloning artifacts. Five of the false mutations involved base transitions. In sequencing of 2472 bases of subject 14's *CYP11B1* gene, 5 apparent misincorporations were detected, from a total of only three PCR reactions (generating 0.9 kb, 1.4 kb and 1.7 kb fragments). This would suggest that the chance of a base being misincorporated in 2.5 kb of sequence obtained from around 5 kb of amplified, cloned DNA is around 0.2%. The Promega *Taq* Polymerase used in this study has a reported misincorporation rate of 1 in 3,500 bp of amplified DNA, which is approximately 0.03% (Promega Ltd., Southampton, UK, personal communication). The higher rate of misincorporation suggested by this study may in part be due to errors in bacterial replication of the cloned DNA. It should be stressed again that results obtained from PCR amplified DNA must be verified using analysis of more than one PCR reaction. Much of the cloning and direct sequencing carried out used DNA from multiple PCR reactions, which meant that sequences obtained were a consensus, rather than the evaluation of a single PCR.

Verification of mutations proved to be extremely time consuming, unless the sequence change affected a restriction endonuclease recognition site. There were problems with PCR reactions failing to amplify DNA, DNA product being lost during purification, and cloning proving difficult or impossible for a particular piece of DNA. Success was often achieved by beginning PCR from different DNA samples from the same individual and attempting to clone different pieces of DNA containing the same desired region of sequence. Other cloning vectors, such as M13 filamentous phage, mp18, and the pGEM-T plasmid (Promega, Southampton, UK), which is designed specifically for cloning PCR products are alternatives to pUC18 that may be of use in further studies.

Not all sequences were clear or readable up to the point of interest, often there were ghost bands, compressions or apparent missing bases. All these artifacts had to be weeded out by repetition or, where possible, restriction enzyme digestion analysis. In short the methods given here require perserverance and a great deal of cross checking.

The library screening of genomic DNA for the CYP17 deficient patient was only partially successful, with the 5' region and 3 exons isolated, but took almost a year to obtain results. Preparation of different libraries using different fragments may have succeeded in isolating the entire gene, but would have been unrealistic in terms of the time available. Fortunately one mutation was isolated from the library screening within the year, and the first PCR amplification of subject 17's *CYP17* exon 4 showed the second mutation present in this patient. Library screening is still useful in obtaining 5' control regions or repetitive regions of DNA not conducive to PCR amplification and in situations where the genomic sequence is not available, but now takes second place to PCR based approaches in terms of speed.

6.6 CYP11B1, CYP17 and hypertension

The rarity of mutation within *CYP17* suggests that mutations at this locus have been selected against. Due to this process polymorphisms that affect CYP17 function may also be rare. Three polymorphisms have been reported, but all involve silent mutations (CAT→CAC, H46H, TCT→TCG, S65S and GAT→GAC, D283D). Differences in the level of expression of CYP17 and the regulation of the 17,20 lyase function may exist, however, and influence hypertension or hyperandrogenism.

In contrast this study has shown that the *CYP11B1* locus is highly polymorphic and there is the potential for splicing abnormalities and enzymes with altered amino acid composition. Two possible polymorphisms described in this thesis, one involving the loss of the intron 1 branch site for splicing and the other the C494F mutation, may reduce the production of cortisol and represent risk factors for hypertension. Mutations with a limited effect on CYP11B1 activity may well have been retained in populations and have more serious effects when combined with other mutations.

In rats CYP11B1 also contributes to mineralocorticoid production (Lauber *et al*, 1989), and a recent study of salt resistant Dahl rats showed them to have 5 amino acid changes within the CYP11B1 enzyme (Cicila *et al*, 1993), which leads to a reduction of the mineralocorticoid 18-hydroxy-11-deoxycorticosterone and salt loss. This being so, there is a possibility that mutation of key amino acids in the human CYP11B1 enzyme could lead to mineralocorticoid production and salt retention. The distribution of polymorphisms within the *CYP11B1* and *CYP11B2* genes therefore warrants further investigation in hypertensive populations.

6.7 Future directions for study of *CYP11B* and *CYP17* genes

Of immediate interest are the effects of the mutations that have been located in this group of 11ß-hydroxylase patients. Mutagenesis and expression studies are necessary to determine whether the amino acid alterations and possible splice site mutations do actually affect the enzyme function. Computer molecular modeling and structure evaluation studies on the CYP11B proteins have not been reported to date, but would be useful in determining which regions of these enzymes are essential for substrate binding, membrane attachment and enzyme stability.

The distribution of the two CYP11B1 polymorphic groups (as defined in section 4.6.2) within hypertensive populations would be of interest. Such a study, together with biochemical assessment of the mineralocorticoid production of the different polymorphic groups, may identify subgroups of hypertensives with steroid associated disease.

Future studies on the variability of CYP17 function are more likely to be concentrated on the control of expression and the regulation of 17,20 lyase activity. The lyase function becomes highly active at adrenarche and declines in old age. The factors controlling this are largely unknown at present. Similarly metabolic defects that affect the level of activity are poorly understood. Exaggerated adrenarche, ovarian hyperandrogenism and other conditions related to androgen excess are widespread and will presumably offer up avenues of research into ancillary factors affecting the level of CYP17 activity.

6.8 Final Conclusion

In conclusion, mutations have been located in *CYP17*, *CYP11B1* and *CYP11B2*, using a combination of gene library screening and PCR-SSCP techniques. In light of the paucity of mutations discovered at these loci this study has greatly increased the range of mutations found and begun to address the question of how mutations have arisen and what affect they could have on enzyme function. Mechanisms leading to these mutations have been proposed and hypotheses put forward to explain the differences in the apparent mutation rate between the three genes. Finally, an attempt has been made to link the mutations found to loss of enzyme function.

The original hypothesis proposed was that deleterious mutations and the increased rate of mutation in the *CYP11B1* gene relative to the *CYP17* gene, were due to gene conversion and non-homologous recombination with the *CYP11B2* gene. This was shown not to be correct. Gene conversion and non-homologous recombination do occur at the *CYP11B* locus, but are not the major cause of pathological mutations.

Deamination of ^{m}CpG or ^{m}C residues appears to have played a large part in generation of mutations in the *CYP11B1* gene, suggesting that it is highly methylated in tissue giving rise to gametes. Misincorporation of nucleotides also accounts for a large number of mutations in *CYP11B1*. Misalignment between repetitive sequences within the gene is a third major cause of mutation, causing duplication and deletion of small regions of DNA.

CYP17 contains a significantly lower number of exonic CpG dinucleotides than *CYP11B1*, *CYP11B2* or *CYP21* and is presumably not highly methylated in the gonads. The main factor causing the apparent low level of deleterious mutation within *CYP17*, however, is presumed to be natural selection removing non-functional *CYP17* alleles from the population. This contrasts with both the *CYP21* and *CYP11B1* gene, where mutations giving rise to hirsutism, accelerated somatic growth and aggressive behaviour may all have conferred advantages on heterozygotes and the populations with which they were associated.

Some questions have been answered, but more have been generated by the work carried out here.

Appendix 1.- PCR Primers

<u>*CYP17* Primers</u>

Primer	Position	Sequence	Tm °C
715S	5' Flank (-336)	AACCACATGGGCTGGAGGGC	66
716S	5' Flank (-18)	GGCAAGATAGACAGCAGTGG	62
947R	5' Ex. 1	CCACTGCTGTCTATCTTGCC	62
952R	3' Ex. 1	TGAAGACCTGAACCAATCCC	60
948R	5' Ex. 2	GGTGTGAGATTCCTACAGCC	62
713S	3' Ex. 2	TCCAGCAGCTCCTGTGGGAT	64
708P	5' Ex. 3	TGCTGATTCATTTCCACCCT	58
709P	3' Ex. 3	AATGTCAGGGTCTACTAGAA	56
949R	5' Ex. 4	GGTGGAGTAGGAACTTCCAG	62
953R	3' Ex. 4	TGTGCCAGGTTCTCTGCTTG	62
950R	5' Ex. 5	TGGCAGGAGTGTCACAGATG	62
714S	5' Ex. 6	TTACACACACTAGTCACCTC	58
954R	3' Ex. 6	CAAGCAGTGAATGCATCAGG	60
951R	5' Ex. 7	TTCCTCTTCCACTCTGGAGC	62
955R	3' Ex. 8	GGCCACATAGGGTGGACAGG	66

In the table above, and that which follows overleaf, oligonucleotide primers were given arbitrary names, e.g. 714S. Primers were chosen that were complementary to intronic sequences immediately 5' or 3' of the exons specified. The primer 716S and those positioned 3' of the exons are the antisense primers, but the sequences above are the actual oligonucleotide sequence, written 5'-3'. Tm values are approximate and were calculated from the Wallace formula, as described in section 2.2.4.

CYP11B1 and *CYP11B2* Primers

Primer	Position	Sequence	Tm °C
072T	5' Flank B1	TTTGAATTCTCGAAGGCAAGGCACCAG	58
073T	3' Ex. 2 B1/B2	GGGGGATCCTGCTCCCAGCTCTCAGCT	58
074T	5' Ex. 3 B1	CCCGAATTCAGAAAATCCCTCCCCCCTA	58
075T	3' Ex. 5 B1	CCCGGATCCGACACGTGGGCGCCGTGTGA	68
076T	5' Ex. 6 B1	CCCGGATCCTGACCCTGCAGCTGTGTCT	60
077T	3' Ex. 9 B1	GAGACGTGATTAGTTGATGGC	62
078T	5' Ex. 8 B1	TACTCTCTGGGTCGCAACCCC	68
079T	3' Ex. 9 B1	GGGGGCACATGCTGGGCCTCA	72
709V	5' Flank B2	TCCTTCATCTACCTTTGGCTGGGG	74
710V	3' Ex. 4 B2	CCGAGACTGCCCCGACACCCAAAT	78
711V	5' Ex. 5 B2	ATTTGGGTGTCGGGGCAGTCTCGG	78
16W	3' Ex. 9 B2	TTGCTATTTGACAAGCCTGGCAAG	70
7W	3' Ex. 1 B1/B2	GAATGGCAGTGCTGAGTGCC	64
8W	5' Ex. 2 B1/B2	TTTGGATTGGGACTGCAGGG	62
9W	5' Ex. 3 B1/B2	CTGCAGGCCGATTCCCCTTG	66
10W	3' Ex. 3 B1/B2	TGGCCACTCCAGGGTCTCTG	56
11W	5' Ex. 4 B1/B2	CCTTGTGCTCAGCAGTGCAT	62
12W	3' Ex. 4 B1/B2	GTGGTGGAGAGGGAGAAATT	60
13W	3' Ex. 6 B1/B2	AGGGCCACAGGGAGGCCTCA	68
14W	5' Ex. 7 B1/B2	TGGATGCCCCCACCTCCAGG	68
15W	3' Ex. 7 B1/B2	AATGACTGGGGAGGGAGGTT	62
368Y	5' Ex. 5 B1/B2	ACTTGGTGCTGGGCTAGCTG	64
369Y	3' Ex. 5 B2	GAACGTGGGTGCCGTGTGGC	68
370Y	5' Ex. 6 B2	GACCCTGCAGACATGGCTTC	64
371Y	5' Ex. 8 B1/B2	CCCTCGAGCTGAGAACCTCC	66
A617	3' Ex. 8 B1/B2	ACATGGTGCAGCAGCAGCAGC	68
A618	5' Ex. 9 B1/B2	CTGTTCCCCCTTCAGCATAAT	62
G3937	5' Ex. 4 B1/B2	GAGCCTGTCTCGCTGGA	56
H1042	5' Ex. 5 B1	CCCGGATCCCAGGAGGACACTGAAGGATGTT	66

In the table above underlined regions of certain primers indicate *Eco*RI or *Bam*HI restriction enzyme sites, GAATTC and GGATCC, which were incorporated during oligonucleotide synthesis to facilitate cloning of PCR fragments subsequent to enzyme digestion.

Appendix 2- The three and one letter codes for amino acids

Shown below are the 20 amino acids. A description of the nature of the side chain and the three and one letter codes for each amino acid is also given.

Amino acid	Type of side chain	3 letter code	1 letter code
Alanine	Hydrophobic	Ala	A
Arginine	Basic	Arg	R
Asparagine	Polar	Asn	N
Aspartic acid	Acidic	Asp	D
Cysteine	Polar/SH group	Cys	C
Glutamic acid	Acidic	Glu	E
Glutamine	Polar	Gln	Q
Glycine	Hydrophobic	Gly	G
Histidine	Basic	His	H
Isoleucine	Hydrophobic	Ile	I
Leucine	Hydrophobic	Leu	L
Lysine	Basic	Lys	K
Methionine	Hydrophobic	Met	M
Phenylalanine	Hydrophobic	Phe	F
Proline	Hydrophobic	Pro	P
Serine	Polar/OH group	Ser	S
Threonine	Polar/OH group	Thr	T
Tryptophan	Polar/aromatic	Trp	W
Tyrosine	Polar aromatic OH group	Tyr	Y
Valine	Hydrophobic	Val	V

N.B. In the text of the thesis mutations altering amino acids are shown as, for example, G267R. The first letter refers to the normal amino acid, the number the position of that amino acid in the protein polypeptide chain and the second letter the amino acid present in the mutant form.

Appendix 3- The eukaryotic genetic code

The triplet codons of eukaryote DNA are shown below. Sixty-one codons represent amino acids. All of the amino acids except tryptophan and methionine are represented by more than one codon. The synonymous codons usually form groups in which the base in the third position has least significance. Three codons, TAA, TAG and TGA, labelled Stop below, cause termination of the peptide chain. The codon sequences are given in the 5'-3' orientation.

First Base		Second Base: T	Second Base: C	Second Base: A	Second Base: G
	T	TTT, TTC } Phe TTA, TTG } Leu	TCT, TCC, TCA, TCG } Ser	TAT, TAC } Tyr TAA, TAG } Stop	TGT, TGC } Cys TGA Stop TGG Trp
	C	CTT, CTC, CTA, CTG } Leu	CCT, CCC, CCA, CCG } Pro	CAT, CAC } His CAA, CAG } Gln	CGT, CGC, CGA, CGG } Arg
	A	ATT, ATC, ATA } Ile ATG Met	ACT, ACC, ACA, ACG } Thr	AAT, AAC } Asn AAA, AAG } Lys	AGT, AGC } Ser AGA, AGG } Arg
	G	GTT, GTC, GTA, GTG } Val	GCT, GCC, GCA, GCG } Ala	GAT, GAC } Asp GAA, GAG } Glu	GGT, GGC, GGA, GGG } Gly

Appendix 4- Published *CYP11B1* gene sequence

The sequence below is taken from the Human Genome databank (Accession numbers: 5' flanking region and exon 1: D10169 D90428 X55765; exon 1 and 2: M32863 J01540; exons 3-5 M32878 J05140; exon 9: M32879 J01540). Lower case letters represent intronic sequence, upper case exonic sequence. The encoded amino acid is given above each codon (see Appendix 2 for 1 letter code for amino acids). The numbering represents the position of each amino acid in the polypeptide chain and the * at codon 504 marks the termination codon. Introns 2 and 8 are incomplete, with approximately 85 and 440 bp uncharacterzied. The position of certain primers used in this study are shown by underlined sequences and labelled with the arbitrary name, as defined in Appendix 1, with arrows indicating sense (→) or antisense (←).

```
     ggctcttggataagataagggccccatccatcttgctcctctcagccctggaggaggagggagagtccttttcccctgtctacgctcatgcacccca
                                                                                                    072T→
atgagtccctgcctccagccctgacctctgccctcggtctctcaggcagatccagggccagttctcccatgacgtgatccctcccgaaggcaaggcaccaggca

                                    Exon 1  M   A   L   R   A   K   A   E   V   C   M   A   V   P   W
agataaaaggattgcagctgaacagggtggagggagcattgga ATG GCA CTC AGG GCA AAG GCA GAG GTG TGC ATG GCA GTG CCC TGG
                                            1                                   10

 L   S   L   Q   R   A   Q   A   L   G   T   R   A   A   R   V   P   R   T   V   L   P   F   E   A   M
CTG TCC CTG CAA AGG GCA CAG GCA CTG GGC ACG AGA GCC GCC CGG GTC CCC AGG ACA GTG CTG CCC TTT GAA GCC ATG
                20                                      30                                      40

 P   R   R   P   G   N   R   W   L   R   L   L   Q   I   W   R   E   Q   G   Y   E   D   L   H   L   E
CCC CGG CGT CCA GGC AAC AGG TGG CTG AGG CTG CTG CAG ATC TGG AGG GAG CAG GGT TAT GAG GAC CTG CAC CTG GAA
                                50                                      60

 V   H   Q   T   F   Q   E   L   G   P   I   F   R
GTA CAC CAG ACC TTC CAG GAA CTG GGG CCC ATT TTC AG gtaaagccctccctggccctcgctggaacacccagtgccctgcccttgctgc
        70                                      80

ccaggaccctgccgggcactcagcactgccattcccagcaggtcccggcactctgcatcctttggaagagggaagatcgagcacgtgctgtctgtgcgctgcag

ggcagggcatgtgcagagcaaatgggagctcggctgcagagagggcaggactcagaggcactgaagttaagaggttccgggcagtcagcaagagggcgtttagc

tgtgaagccgctaatccaggagaggggagggtggacaggagacactttggattgggactgcagggtggggccagcagggactagaccccgtccagcagggcct

           Exon 2    Y   D   L   G   G   A   G   M   V   C   V   M   L   P   E   D   V   E   K   L   Q
cctgcttggccccacag G TAC GAT TTG GGA GGA GCA GGC ATG GTG TGT GTG ATG CTG CCG GAG GAC GTG GAG AAG CTG CAA
                                                        90                                      100

 Q   V   D   S   L   H   P   H   R   M   S   L   E   P   W   V   A   Y   R   Q   H   R   G   H   K   C
CAG GTG GAC AGC CTG CAT CCC CAC AGG ATG AGC CTG GAG CCC TGG GTG GCC TAC AGA CAA CAT CGT GGG CAC AAA TGT
                                110                                     120

 G   V   F   L   L
GGC GTG TTC TTG CT gtaagcggcgagctgagagctgggagcagggtgggcagcctgggtgtagggggaggcgagagaggcagga .... (85 bp)
        130                       ← 073T

aaagcttgacaacaggggtcagttcctttcttgcagaaaatccctcccccctactacagggagggcccgcatgggtgaggtggtgccagacttggggcgccagg
                                             074T →
tcccgggaatgacctcagttaccctgtcagcacctgtgggcagaagctaccatctcatccctgcttagacctgagtggcctttgtccagcacctggaggccgtc
```

```
tgagaaaaggctgcagctcgaacacaaacaggcagcttctaccagggcccccagtcagctccctgcaggccgattccccttgggacaaggaggatgggatacgg

                                                      Exon 3  N   G   P   E   W   R   F   N   R   L   R
gtcagggcctgtgttttgctggggcggcctcacaagctctgccctggcctctgtag G AAT GGG CCT GAA TGG CGC TTC AAC CGA TTG CGG
                                                                                          140

 L   N   P   E   V   L   S   P   N   A   V   Q   R   F   L   P   M   V   D   A   V   A   R   D   F   S
CTG AAT CCA GAA GTG CTG TCG CCC AAC GCT GTG CAG AGG TTC CTC CCG ATG GTG GAT GCA GTG GCC AGG GAC TTC TCC
                        150                                     160

 Q   A   L   K   K   K   V   L   Q   N   A   R   G   S   L   T   L   D   V   Q   P   S   I   F   H   Y
CAG GCC CTG AAG AAG AAG GTG CTG CAG AAC GCC CGG GGG AGC CTG ACC CTG GAC GTC CAG CCC AGC ATC TTC CAC TAC
170                                     180                                     190

 T   I   E
ACC ATA GAA G gtgtgggccacatgggttgatccagcctcagagaccctggagtggccagggacggggatgggggactgaagggagtgtggggaggcagcc

                                             Exon 4 A   S   N   L   A   L   F   G   E   R   L   G   L   V   G
aggaggcccggttcccttgtgctcagcagtgcatcctccccgcag CC AGC AAC TTG GCT CTT TTT GGA GAG CGG CTG GGC CTG GTT GGC
                                                 200                                     210

 H   S   P   S   S   A   S   L   N   F   L   H   A   L   E   V   M   F   K   S   T   V   Q   L   M   F
CAC AGC CCC AGT TCT GCC AGC CTG AAC TTC CTC CAT GCC CTG GAG GTC ATG TTC AAA TCC ACC GTC CAG CTC ATG TTC
                        220                                     230

 M   P   R   S   L   S   R   W   T   S   P   K   V   W   K   E   H   F   E   A   W   D   C   I   F   Q
ATG CCC AGG AGC CTG TCT CGC TGG ACC AGC CCC AAG GTG TGG AAG GAG CAC TTT GAG GCC TGG GAC TGC ATC TTC CAG
240                                     250                                     260

 Y
TAC G gtgaggccagggacccgggcagtgctatggggaaggacaccatcgggcccccaatttctccctctccaccacccagtggggaatggaggccacaggg

aggggtcggggattcctcaccgtcctgccagggagattggtgtgaggctggggctgggctgggctgatccggagaatttgggatgagagcagggagacttggt

                                                                                  Exon 5 G   D   N   C   I
gctgggctagctggcaggaggaggacactgaaggatgtttcccagcaccaaagtctgagggctgcctcccgctccccggatag GC GAC AAC TGT ATC
                                                                                                270

 Q   K   I   Y   Q   E   L   A   F   S   R   P   Q   Q   Y   T   S   I   V   A   E   L   L   L   N   A
CAG AAA ATC TAT CAG GAA CTG GCC TTC AGC CGC CCT CAA CAG TAC ACC AGC ATC GTG GCG GAG CTC CTG TTG AAT GCG
                                280                                     290

 E   L   S   P   D   A   I   K   A   N   S   M   E   L   T   A   G   S   V   D   T
GAA CTG TCG CCA GAT GCC ATC AAG GCC AAC TCT ATG GAA CTC ACT GCA GGG AGC GTG GAC ACG gtcaggccggcaaccagc
        300                                     310

cccacccagagagggtgatgccaagcctgcctcccaggcactgcctgccaatgtcacacggcgcccacgtgtcccatgcccaggctatgggccccacatttctt
                                                          ← 075T
acttgggattgtgatgtgataaacacgtttgcaggttgccatggttggaatggggggttcctttccttctgtggaggactcagggaaacggggtttggatgggc

attaggatttgaagtcttgggctctgtcgtgctcagggtatgcatgtctgcacccctcacagggaggttgtcctgggaggggtgtcccgggggctgagtcctcc

          076T →        Exon 6 T   V   F   P   L   L   M   T   L   F   E   L   A   R   N   P
tgtgcaaggtctgaccctgcagctgtgtctcctgcag ACG GTG TTT CCC TTG CTG ATG ACG CTC TTT GAG CTG GCT CGG AAC CCC
                                          320                                     330

 N   V   Q   Q   A   L   R   Q   E   S   L   A   A   A   A   S   I   S   E   H   P   Q   K   A   T   T
AAC GTG CAG CAG GCC CTG CGC CAG GAG AGC CTG GCC GCC GCA GCC AGC ATC AGT GAA CAT CCC CAG AAG GCA ACC ACC
                    340                                     350                                     360

 E   L   P   L   L   R   A   A   L   K   E   T   L   R
GAG CTG CCC TTG CTG CGT GCG GCC CTC AAG GAG ACC TTG CG gtgggtgctggctgaggcctccctgtggccctggccctgctggagagt
                                    370

cagcccccactgggtggttgcagacagaatctgggctataaacacctacccagcagccatcctgactgctctctcgcgtcaaggacagggagctcttcttcctc

tggaatccctcttcaacgccctggggattaacgtgggggcatgtccttctgcgctcggggctgcttaagttaggggaggtttggccgggctcagcaggtgcaag

gaagcacttcctacacctgggcttcccatggatctgggacctctgcggggtcttcggtaggaagggtgcagagagcacaggaaccccatcccagctgagaccct

                      Exon 7  L   Y   P   V   G   L   F   L   E   R   V   A   S   S   D   L   V   L   Q
ttctatggatgcccccacctccag G CTC TAC CCT GTG GGT CTG TTT CTG GAG CGA GTG GCG AGC TCA GAC TTG GTG CTT CAG
                                              380                                     390

 N   Y   H   I   P   A   G
AAC TAC CAC ATC CCA GCT GGG gtgagtgagccccacaccctcgagctgagaacctccctccccagtcattccctgatccccgctctgctccgtccgc
                        400            078T →
```

```
Exon 8
   T   L   V   R   V   F   L   Y   S   L   G   R   N   P   A   L   F   P   R   P   E   R   Y   N   P
ag ACA TTG GTG CGC GTG TTC CTC TAC TCT CTG GGT CGC AAC CCC GCC TTG TTC CCG AGG CCT GAG CGC TAT AAC CCC
                                  410                                     420

 Q   R   W   L   D   I   R   G   S   G   R   N   F   Y   H   V   P   F   G   F   G   M   R   Q   C   L
CAG CGC TGG CTA GAC ATC AGG GGC TCC GGC AGG AAC TTC TAC CAC GTG CCC TTT GGC TTT GGC ATG CGC CAG TGC CTT
                430                                     440                                     450

 G   R   R   L   A   E   A   E   M   L   L   L   L   H   H
GGG CGG CGC CTG GCA GAG GCA GAG ATG CTG CTG CTG CTG CAC CAT gtgagcaggcccggg.............(440 bp)
                                460

gaattctgggcctgggctgtaaggtggggctggtcaggaatgaaacaggttggaggccaggctgctgttccccttcagcataatctctgcaactttgagggtc

                                                            Exon 9  V   L   K   H   L   Q   V   E   T
tgagaaggctgcaccacgtcgatgggctgcggaccaagccagatggaaacccggcttctgtcctag GTG CTG AAA CAC CTC CAG GTG GAG ACA
                                                                                   470

 L   T   Q   E   D   I   K   M   V   Y   S   F   I   L   R   P   S   M   C   P   L   L   T   F   R   A
CTA ACC CAA GAG GAC ATA AAG ATG GTC TAC AGC TTC ATA TTG AGG CCC AGC ATG TGC CCC CTC CTC ACC TTC AGA GCC
                480                                     490     ← 079T                          500

 I   N   *
ATC AAC TAA tcacgtctctgcacccagggtcccagcctggcaccagcctccctttctgcctgaccccaggccacccctcttctcccacatgcacagcttcc
      ← 077T

tgagtcacccctctgtctaaccagccccagcacaaatggaactcccgagggcctctaggaccagggtttgccaggctaagcagcaatgccagggcacagctgggga

agatcttgctgaccttgtccccagccccacctggccctttctccagcaagcactgtcctctggcagtttgcccccatccctcccagtgctggctccaggctcctcg

tgtggccatgcaagggtgctgtggttttgtcccttgccttcctgcctctagtctcacatgtccctgttcctcttcccctgccagggcccctgcgcagactgtcaga

gtcattaagcgggatcccagcatctagagtccagtcaagttccctcctgcagcctgccccctaggcagctcgagcatgccctgagctctctgaaagttgtcgccct

ggaatagggtcctgcagggtagaataaaaggcccctgtggtcacttgtcctgacatccccattttcaagtgatacaactgagtctcgagggacgtgtgttcccc

agctgatcgtgtcagcctcatgcccctggcctcatctttcatggaccaggccttgttccaggagtgggcgttgggtcctctgcttcctgtgctgtcccctgggg

aaggtcccaaggatgctgtcaggagatggaagagtcatgtggggtgggaacctggggtgtggttccagaaatgtttttggcaacaggagagacaggattgggcc

aacaaggactcagacgagttttattgactattctctgaca
```

Appendix 5- Published *CYP11B2* gene sequence

Below is shown the normal *CYP11B2* sequence (Accession numbers: 5' flanking region and exon 1: D10170 D90429; exons 1 and 2:M32864 J05140; exons 3-8: M32880 J05140; exon 9 and 3' flanking: M32881 J05140). Letters above the sequence represent amino acids encoded in *CYP11B2*. Where two letters are shown (eg. R/Q) the second letter represents an amino acid that differs in *CYP11B1*. Letters below the sequence represent exonic base differences in *CYP11B1*. The numbers below the sequence indicate amino acid positions and underlined bases show primers used for PCR amplification, which are named as defined in Appendix 1.

```
gaattctgca tcctgtgaaa ttatccttca aaagtgaaca taaatatttt ctcaggtaaa taaaaattga ggggattcgc tgccaataga
actgacttgc cagaaatgtt ttttaaaagt tctgcagaga gaaagaaaat gatacaggtc agcaaccctg agctacataa agaaaggaag
agcatttcag aaggaatcag taaagagaaa atgaagtctt ttattttttc ttaatcttaa ttgatctaag agtttgctaa aacaaaacaa
caacaataaa aataggccgg gcgcggtgac tcaccacacc tgtaatccca gcactttggg aggccgaggt gggcagatca cctgaggtca
ggagttcgag actagcctgg ccaacacggt gaaaccccgt ctctactaaa aatacaaaaa ttagctgggc gtggtggtgg cacatgcctg
taatcccagc tacttgggag gctgaggcgg gagaattgct tgaacccggg agacagaggt cgcagtgagc cgagatcaca ccattgcact
ctagcctggg cgacagagtg agactctgtc tcaaaaataa ataaataaat aaataaataa ataataaaaa taaataaata aaagccagaa
agtgtatttg atgatcatag ttatgtatat gtgaaatgaa ggacagcaat gatgcaaggg atgggtgagt ggaattaaaa atatcttatt
atttatttat tttgagatgg agtcttgctt tgctgcccag gttggagtgc agtgggatga tctcaactca ctgcaacctc cgcctccttg
attcaagcat tcatcttgac tcagcctgct gagaagccga gattacaggc atgcgccacc acacctggct aattttgtat ttttagtaga
gacagggttt tgccatgttg gccaggctgg tctcgaactc ctgacctcag gtgatccacc tgcatcagcc tcccaaagtg ctgggatgac
agacatgagc cactatgccc agcctaagaa tatctgatga ttataaagtg cttgcattac ctctgaagct gtatagtgtt atatgaaggt
ggagttggag agatgagttt taagcgtata ttgcaaactc tagggcaacc actaaagaag tgagacccag cctctagaaa aaaaaaaaaa
aaaggaaatt agctatcaag ccacgaaaag aaatggagga accttaaacg catattacta actgagatac gtcactttga aaaggctaca
aacggtgtca ttccaactat acaacatttt ggaaaaggcc aaagcatggt gatgataaaa aagatcggag atgtcaggga ctggggcagg
agggatgagc aggcagagca caggttttct tttcctcttt ttaagacagt gaaaatactc ctaggatcct gcaaggaggg atacaaatta
catacatttg tcaaaaccca cagcatgttg accaccagga ggagacccca tgtgactcca ggaccctggt tgataacaac gtatcgagat
tcctcacatg gaaccagtgc gctcctgtgg tggagggtgt acctgtgtca gggcaggggg tacgtggaca ttttctgcag tttttgatca
attttgcaat gaactaaatc tgtggtataa aaataaagtc tattaaaaga atccaaggct ccctctcatc tcacgataag ataaagtccc
catccatttt actcctctca gccctggaga aaggagaggc caggtcccac caccttccac cagcatggac ccccagtcca gaccccacgc
                                                                     709V →
cttttctcag catcctcaga ccagcaggac ttgcagcaat ggggaattag gcacctgact tctccttcat ctacctttgg ctggggcctc
cagccttgac cttcgctctg agagtctcag gcaggtccag aggccagttc tcccatgacg tgatatgttt ccagagcagg ttcctgggtg
agataaaagg atttgggctg aacagggtgg agggagcatt gga

                                                        Exon 1 M   A   L   R   A   K   A   E   V   C  V/M  A  A/V  P   W
                                                               ATG GCA CTC AGG GCA AAG GCA GAG GTG TGC GTG GCA GCG CCC TGG
                                                                                                        A       T
                                                                1                                   10

L  C/S  L   Q   R   A  R/Q  A   L   G   T   R   A   A   R  A/V  P   R   T   V   L   P   F   E   A   M
CTG TGC CTG CAA AGG GCA CGG GCA CTG GGC ACT AGA GCC GCT CGG GCC CCT AGG ACG GTG CTG CCG TTT GAA GCC ATG
     C                  A               G           C       T   C       A           C
                20                                      30                                      40

P  Q/R H/R  P   G   N   R   W   L   R   L   L   Q  M/I  W   R   E   Q   G   Y   E  H/D  L   H   L   E
CCC CAG CAT CCA GGC AAC AGG TGG CTG AGG CTG CTG CAG ATG TGG AGG GAG CAG GGT TAT GAG CAC CTG CAC CTG GAG
     G   G                                           C                               G               A
                                    50                                      60
```

```
M/V H   Q   T   F   Q   E   L   G   P   I   F   R
ATG CAC CAG ACC TTC CAG GAG CTG GGG CCC ATT TTC AG gtaaagccctccctggccctcgctgggaacacccagatccctccccctgctgc
G A                      A
     70                                       80
                      ← 07W
ccaggacactgccaggcactcagcactgccattcccagcaggtcccggcactctgcatcctttggaggatggggaaggagtgcagcacatgctggtctgtggtg

ctgccagggcaggggatagtgcagagaaaccccagctcactgcagagagggcaggactcagaagcactaaagttgaaaggttccagggagccagcaggagggc
                                                                                          08W →
tttagctgtgaagccgctaatccaggagcagggagggtggacaggagacactttggattgggactgcagggtggggccacgagggacatgaccccgtccagcag

                  Exon 2    Y  N/D  L   G   G  P/A R/G  M   V   C   V   M   L   P   E   D   V   E   K
ggcctcctgcttggccccacag G TAC AAC TTG GGA GGA CCA CGC ATG GTG TGT GTG ATG CTG CCG GAG GAT GTG GAG AAG
                             G T         G   G                                       C
                                                          90

 L   Q   Q   V   D   S   L   H   P  C/H  R   M  I/S  L   E   P   W   V   A  I/Y  R   Q   H   R   G   H
CTG CAA CAG GTG GAC AGC CTG CAT CCC TGC AGG ATG ATC CTG GAG CCC TGG GTG GCC ATC AGA CAA CAT CGT GGG CAC
                                    CAC         G                       TAC
100                                     110                                 120
                                                ← 073T
 K   C   G   V   F   L   L
AAA TGT GGC GTG TTC TTG TT    gtaagcggcgagttgggagctgagagctgggagcagggtgggcagcctgggtgtagggggggaggcgagagaggtag
                 C
                 130

gacccaaaagcacatctgccctgggcccctgtggtgggcagtgagggtgagcacccggcccagaggacggccatcctgtggggtcgcgtctgcactgtgggttg

gggaagcagggcggtggtggagaaatgggcaggggcacctctgcagagaagacgcagagcaatgagcccttctgtgtagtgagaacccgctctgcaccaacctc

ggcggctgctttctcttgcggtctggggactctccttcccataggtcagaaaactgaggccctgagaaggggacttccactggcccaggtcacaggctgagtac

tgagcctggtgttcgccggggccacagcctccctcagggcgctcagggtccctgcag....

..aacaggggtcacctcctttcttggagaaaagccctaccctgttactacagggagggcctgcatgggtgaggtggtgccagacttgggtcgccaggtccca

ggaatgacctcagttaccctgtcagcacctgtgggcagaagctacagtctcatccctgcttagacctgagcggcctttgcccagcacctggaggtcgctctg

                                                                                 09W →
agaaaaggtctgcagctcgaacacaaacaggcagcttctaccagggccccagtcactcctgcaggccgattccccttgggtacaaggaggatgggatacggg

                                                        Exon 3   N   G   P   E   W   R   F   N   R   L   R
gtcagggcctgtgtcttgctggggcggcctcacaagctctgccctggcctctgtag G AAT GGG CCT GAA TGG CGC TTC AAC CGA TTG CGG

                                                                                         140

 L   N   P  D/E  V   L   S   P  K/N  A   V   Q   R   F   L   P   M   V   D   A   V   A   R   D   F   S
CTG AAC CCA GAT GTG CTG TCG CCC AAG GCC GTG CAG AGG TTC CTC CCG ATG GTG GAT GCA GTG GCC AGG GAC TTT TCC
      T       A               C   T                                                             C
                         150                                     160

 Q   A   L  R/K  K   K   V   L   Q   N   A   R   G   S   L   T   L   D   V   Q   P   S   I   F   H   Y
CAG GCC CTG AGG AAG AAG GTG CTG CAG AAC GCC CGG GGG AGC CTG ACC CTG GAC GTC CAG CCC AGC ATC TTC CAC TAC
             A
170                                     180                                     190

 T   I   E                                     ← 10W
ACC ATA GAA G gtgtgggccatgcgggaaggtccagccccagagaccctggagtggccagggatggggatggaggactgaagggagtgtggggaggcagcc

                  11W →                  Exon 4 A   S   N   L   A   L   F   G   E   R   L   G   L   V   G
aggaggtccggggctgccttgtgctcagcagtgcatcctccccgcag CC AGC AAC TTA GCT CTT TTT GGA GAG CGG CTG GGC CTG GTT GGC
                                                             G
                                                     200                                     210


 H   S   P   S   S   A   S   L   N   F   L   H   A   L   E   V   M   F   K   S   T   V   Q   L   M   F
CAC AGC CCC AGT TCT GCC AGC CTG AAC TTC CTC CAT GCC CTG GAG GTC ATG TTC AAA TCC ACC GTC CAG CTC ATG TTC
                        220                                     230

 M   P   R   S   L   S   R   W  I/T  S   P   K   V   W   K   E   H   F   E   A   W   D   C   I   F   Q
ATG CCC AGG AGC CTG TCT CGC TGG ATC AGC CCC AAG GTG TGG AAG GAG CAC TTT GAG GCC TGG GAC TGC ATC TTC CAG
                                 C
240                                     250                                     260

 Y                                                                     ← 12W
TAC G gtgaggccagggacccgggcagtgctatggggaagggacaccatgggggcccaatttctccctctccaccacccagtggggaatggaggccacaggg
```

```
                                                                                                  ← 710/711V→
aggggtcggggattcctcaccttcctgccggggagattggtgcgaggctggggctgggctgggctgatccggagaatttgggatgagagcagggagatttggg

 ←710/711V →             H1042 →                                                       Exon 5 G   D   N   C   I
tgtcggggcagtctcggcaggaggaggacactgaaggatgcttcccagcaccaagatctagggctgtcccctgctccctgtacag GT GAC AAC TGT ATC
                                                                                               C
                                                                                                      270

 Q   K   I   Y   Q   E   L   A   F  N/S  R   P   Q  H/Q  Y   T  G/S  I   V   A   E   L   L   L  R/N  A
CAG AAA ATC TAC CAG GAA CTG GCC TTC AAC CGC CCT CAA CAC TAC ACA GGC ATC GTG GCA GAG CTC CTG TTG AAG GCG
             T                       G                   G       C A           G                   T
                                 280                                       290

 E   L   S  L/P E/D  A   I   K   A   N   S   M   E   L   T   A   G   S   V   D   T
GAA CTG TCA CTA GAA GCC ATC AAG GCC AAC TCT ATG GAA CTC ACT GCA GGG AGC GTG GAC ACG gtcaggccagcaaccagc
          G   C     T
        300                                     310

                                                                  ←369Y
cccacccagagagggtgatgccaagccctgcctcccagcactgcctgccaatgccacacggcacccacgttccccatccccaggctacaggccccacatttct

gttgccctcagccttcccctcctttgttaagggatgagatttgcaggggaggggaaatgtgagctcccctcacatgagactgagtttgcagttacctgtgt

ggggatccatgctccaggctggaagaaagttggatgaggccctggacacacagcagctctgtccccactggaaagctctgggtgtacaaggagaaggagggtt

gagaggcagctggaggactccactgggcacccttcccagtgtgcccggtcaccttgggccagaaatgtacatgcatgggagggcagggttgtggggaaggcag

cagcacgggctccagccagtgcagaggggcctgtgggtgcacagtggggagaactcaatggaagcagagggagctggggctccagaactcccaggatgatgct

gaggtctggccccttttctaaggtggctgtgagaacccgcctgaagaggctgcaggggacctgggccttggtggagatgggggtcagctttgcgtgaagaag

tcagggaatctggcccaagtggtcatcaaggtttcagatccggcgtcccagggctctgtcgtgctcagggcatggatgtctccacccctcagagggaggttgt

cctggctggggtgtcccgggggctgagtcctcc

              370Y →           Exon 6  T  A/V  F   P   L   L   M   T   L   F   E   L   A   R   N   P
tgtgcaaggtcagaccctgcagacatggcttctgtag ACA GCG TTT CCG TTG CTG ATG ACG CTC TTT GAG CTG GCT CGG AAC CCC
                                      G   T       C
                                         320                                     330

D/N  V   Q   Q  I/A  L   R  K/Q  E   S   L   A   A   A   A   S   I   S   E   H   P   Q   K   A   T   T
GAC GTG CAG CAG ATC CTG CGC AAG GAG AGC CTG GCC GCC GCA GCC AGC ATC AGT GAA CAT CCC CAG AAG GCA ACC ACC
A               GC          C
                    340                                       350                                     360

 E   L   P   L   L   R   A   A   L   K   E   T   L   R                   ← 13W
GAG CTG CCC TTG CTG CGG GCG GCC CTC AAG GAG ACC TTG AG gtgggtgctggatgaggcctccctgtggccctggcccccctgctggaga
                     T                           C
                                    370

gcagcccccactgggtggtggcagacagaatctggggctgataaacagcgtcacccagcagcccattcccctgcacctgctcttcctcccctcaaggtctgg

gagctcttcttcctctgaatccctcttcaacaccctggggattaacgtggggcatgtccttctgcgcttggggcttctcaagttaggggaggtttggctgggc

cagcaggtgcaaggaagcacttcgtcacgacctgggcttcccatgggccagggagctgtgcggggtcttcggtaggaagggtgcagagagcacagggagcccc

                          14W →
atccagctgaggaccctttctgtggatgcccccacctccag

                Exon 7   L   Y   P   V   G   L   F   L   E   R   V  V/A  S   S   D   L   V   L   Q
                       G CTC TAC CCT GTG GGT CTG TTT TTG GAG CGA GTG GTG AGC TCA GAC TTG GTG CTT CAG
                                                     C               C
                                             380                                     390

 N   Y   H   I   P   A   G                        371Y →      ←  15W
AAC TAC CAC ATC CCA GCT GGG gtgagtgagcccccacaccctcgagctgagaacctccctccccagtcattccctgatccctgctctgcaccg
                        400

Exon 8   T   L   V  Q/R  V   F   L   Y   S   L   G   R   N  A/P  A   L   F   P   R   P   E   R   Y
tccgcag ACA TTG GTA CAG GTT TTC CTC TAC TCG CTG GGT CGC AAT GCC GCC TTG TTC CCG AGG CCT GAG CGG TAT
                    G   GC   G                T                C C                                 C
                                        410                                     420
```

```
 N   P   Q   R   W   L   D   I   R   G   S   G   R   N  L/F H/Y  H   V   P   F   G   F   G   M   R   Q
AAT CCC CAG CGC TGG CTA GAC ATC AGG GGC TCC GGC AGG AAC TTG CAC CAC GTG CCC TTT GGC TTT GGC ATG CGC CAG
 C                                                        C T
                        430                                     440

                                               ← A617
 C   L   G   R   R   L   A   E  A/V  E   M   L   L   L   L   H   H
TGC CTC GGG CGG CGC CTG GCA GAG GCA GAG ATG CTG CTG CTG CTG CAC CAC gtaagcaggcctggg.............
      T                          T                                T
450                                     460

      A618 →
..ccccttcagcataattgttgcacctgggacgatgggaggaagctgccccaggtccatgggctactgaccaggcctgatggaaacccagcctctgtcctag


                             Exon 9  V   L   K   H  F/L L/Q  V   E   T   L   T   Q
                                    GTG CTG AAG CGC TTC CTG GTG GAG ACA CTA ACT CAA
                                              A  A  C    A                     C
                                                470

 E   D   I   K   M   V   Y   S   F   I   L   R   P  G/S T/M S/C  P   L   L   T   F   R   A   I   N   *
GAG GAC ATA AAG ATG GTC TAC AGC TTC ATA TTG AGG CCT GGC ACG TCC CCC CTC CTC ACT TTC AGA GCG ATT AAC TAG
                                                  C A    T   G                  C           C   C       A
    480                                     490                                     500        503

tcttgcatctgcacccagggtcccagcctggccaccagcttccctctgcctgaccccaggccacctgtcttctcccacgtgcacagcttcctgagtcacc
                                                                   ← 16W
cctctgtccagccagctcctgcacaaatggaactccccagggcctccaggactggggcttgccaggcttgtcaaatagcaaggccagcgcacagctggagcgat

cttgctgcaggcctgccttgtccccagccccacctggccccttctccagcaagcagtgccctctggacacttgactctactcctcccagcgctggctccaggct

cctcatgaggccatgcaagggtgctgtgattttgtcccttgccttcctgcctagtctcacatgtccctgtccctctcgccctggccagggcctctgtgcagaca

gtgtcagagtcattaagcgggatcccagcatctcagagtccagtcaagttccctcctgcagcctgaccccaggcagctcgagcatgccctgagctctctgaaag

ttgtcacccagaaatacgatcctgcagggtagactaaaaaggcccctgtggtcacttatactgacacattttaagtgatacaactgagtctcgaggggcgtgtg

ttccccagctgatcatgtcagcctcatgccccaggcctcgtctttcatggaccaggtcttgttcaagcagcgagtgttgggtcctctgcttcctgagctgtccc

ctggaaaaggtcccgaggatgctgtcaggagatggaagagtcatgtggggtgggaacctggggtgtggttccagaaatgttttggggcaacaggagacaggatt

gggccaacaaggactcagatgagtttattgactcattcctctggaagatacgcagc.
```

Appendix 6- Exonic differences between *CYP11B1* and *CYP11B2* genes

Given below are the exonic sequence differences between *CYP11B1* and *CYP11B2* genes; *CYP11B2* sequence is written first, followed by the corresponding sequence in *CYP11B1*. Possible explanations for the differences are given, assuming that the points of difference are caused by mutation of a common ancestral gene sequence. For discussion of the distribution of possible mutations see section 6.2..

Codon	*CYP11B2/ CYP11B1* Sequence: Possible mutation of ancestral gene
11	TGC GTG/ TGC ATG: ^{m}CpG deamination B2→B1
13	GCG/ GTG: ^{m}CpG deamination B2→B1
17	TGC/ TCC: Misincorporation G→C/ C→G
22	CGG/ CAG: ^{m}CpG deamination B2→B1
26	ACT/ ACG: Misincorporation T→G/ G→T
29	GCT/ GCC: Misincorporation/ deamination ^{m}C
31	GCC/ GTC: Misincorporation/ deamination ^{m}C
32	CCT/ CCC: Misincorporation/ deamination ^{m}C
34	ACG/ ACA: ^{m}CpG deamination B2→B1
37	CCG/ CCC: Misincorporation G→C/ C→G
43	CAG/ CGG: ^{m}CpG deamination B1→B2
44	CAT/ CGT: ^{m}CpG deamination B1→B2
55	ATG/ ATC: Misincorporation G→C/ C→G
63	CAC/ GAC: Misincorporation C→G/ G→C
67	GAG/ GAA: Misincorporation G→A/ A→G
68	ATG/ GTA: Double misincorporation A→G/ G→A
74	GAG/ GAA: Misincorporation G→A/ A→G
82	TAC AAC/ TAC GAC: ^{m}CpG deamination B1→B2 TAC GAT deamination of ^{m}C

86	CCA/ GCA: Misincorporation C→G/ G→C
87	CGC/ GGC: Misincorporation C→G/ G→C
96	GAT/ GAC: Misincorporation/ deamination of ^{m}C
109	TGC/ CAC: Suggests ancestral gene CGC: ^{m}CpG deamination B2→B1 + B1→B2
112	ATC/ AGC: Misincorporation T→G/ G→T
119	ATC/ TAC: Double misincoporation A→T/ T→A
132	TT/ CT: Misincorporation/ deamination of ^{m}C
145	AAC/ AAT: Misincorporation/ deamination of ^{m}C
147	GAT/ GAA: Misincorporation T→A/ A→T
152	AAG/ AAC: Misincorporation G→C/ C→G
153	GCC GTG/ GCT GTG: ^{m}CpG deamination B2→B1
168	TTT/ TTC: Misincorporation/ deamination of ^{m}C
173	AGG/ AAG: Misincorporation G→A/ A→G
202	TTA/ TTG: Misincorporation A→G/ G→A
248	ATC/ ACC: Misincorporation/ deamination of ^{m}C
267	GC GAC/ GT GAC: ^{m}CpG deamination B2→B1
275	TAC/ TAT: Misincorporation/ deamination of ^{m}C
281	AAC/ AGC: Misincorporation A→G/ G→A
285	CAC/ CAG: Misincorporation C→G/ G→C
287	ACA/ ACC: Misincorporation A→C/ C→A
288	GGC/ AGC: Misincorporation G→A/ A→G
291	GCA/ GCG: ^{m}CpG deamination B1→B2
296	AAT/ AAG: Misincorporation T→G/ G→T
300	TCA/ TCG: ^{m}CpG deamination B1→B2
301	CTA/ CCA: Misincorporation/ deamination of ^{m}C
302	GAA/ GAT: Misincorporation A→T/ T→A
319	ACA/ ACG: ^{m}CpG deamination B1→B2
320	GCG/ GTG: ^{m}CpG deamination B2→B1
322	CCG/ CCC: Misincorporation G→C/ C→G
335	CCC GAC/ CCC AAC: ^{m}CpG deamination B2→B1
339	ATC/ GCC: Double misincorporation/ deamination of ^{m}C

342	AAG/ CAG: Misincorporation A→C/ C→A
366	CGG/ CGT: Misincorporation G→T/ T→G
374	AG/ CG: Misincorporation A→C/ C→A
382	TTG/ CTG: Misincorporation/ deamination of ^{m}C
386	GTG/ GCG: ^{m}CpG deamination B1→B2
403	GTA/ GTG: Misincorporation A→G/ G→A
404	CAG/ CGC: Suggests ancestral gene CGG: ^{m}CpG deamination B1→B2+ Misincorporation G→C
405	GTT/ GTG: Misincorporation T→G/ G→T
409	TCG/ TCT: Misincorporation G→T/ T→G
413	AAT GCC/ AAC CCC: Suggests ancestral gene AAC GCC: ^{m}CpG deamination B1→B2
414	GCC/ CCC: Misincorporation G→C/ C→G
422	CGG/ CGC: Misincorporation G→C/ C→G
424	AAT/ AAC: Misincorporation/ deamination of ^{m}C
438	TTG/ TTC: Misincorporation G→C/ C→G
439	CAC/ TAC: Misincorporation/ deamination of ^{m}C
451	CTC/ CTT: Misincorporation/ deamination of ^{m}C
458	GCA/ GTA: Misincorporation/ deamination of ^{m}C
466	CAT/ CAC: ^{m}CpG deamination B1→B2
469	AAG/ AAA: Misincorporation G→A/ A→G
470	CGC/ CAC: ^{m}CpG deamination B2→B1
471	TTC/ CTC: Misincorporation/ deamination of ^{m}C
472	CTG/ CAG: Misincorporation T→A/ A→T
477	ACT/ ACC: Misincorporation/ deamination of ^{m}C
491 492	CCT GGC/ CCC AGC: Suggests ancestral sequence CCC GGC: ^{m}CpG deamination B1→B2 + B2→B1
493	ACG/ ATG: ^{m}CpG deamination B2→B1
494	TCC/ TGC: Misincorporation C→G/ G→C
498	ACT/ ACC: Misincorporation/ deamination of ^{m}C
501	GCG/ GCC: Misincorporation G→C/ C→G
502	ATT/ ATC: Misincorporation/ deamination ^{m}C
504	TAG/ TAA: Misincorporation G→A/ A→G

Appendix 7- *CYP17* exonic CpG dinucleotides

The table below lists the 38 exonic CpG dinucleotides in *CYP17*, with their position and the result if deamination of ^{m}CpG occurred at those points. Codons marked with * denote points at which such mutations have been reported.

No.	Codon	Sequence	Result if deamination of ^{m}CpG occurred
1	4	CTC GTG	CTT/ ATG L4L/ V5M
2	67	CGT	TGT/ CAT R67C/ H
3	76	GTC GGC	GTT/ AGC V77V/ G78S
4	96	CGG	TGG/ CAG R96W/ Q
5	105	GCG	GCA/ GTG A105A/ V
6	109	CGT	TGT/ CAT R109C/ H
7	112	ATC GCC	ATT/ ACC I112I/ A113T
8	114	TTC GCT	TTT/ ACT F114F/ A115T
9	118	GGC GCA	GGT/ ACA G118G/ A119T
10	125	CGA	TGA/ CAA R125X/ Q
11	128	GCG	GCA/ GTG A128A/ V
12	138	GGC GAT	GGT/ AAT G138G/ D139N
13	161	AAC GGA	AAT/ AGA N161N/ G162R
14	172	TTC GTG	TTT/ ATG F172F/ V173M
15	174	GCG	GCA/ GTG A174A/ V
16	239*	CGA	TGA*/ CAA R239X/ Q
17	255	CGG	TGG/ CAG R255W/ Q
18	300	GGC GTG	GGT/ ATG G300G/ V301M
19	329	TAC GAG	TAT/ AAG Y329Y/ E330K

20	340	CGC	TGC/ CAC R340C/ H
21	347	CGT	TGT/ CAT R347C/ H
22	349	CGT	TGT/ CAT R349C/ H
23	358	CGA	TGA/ CAA R358X/ Q
24	362	CGC	TGC/ CAC R362C/ H
25	365	CCC GTG	CCT/ ATG P365P/ V366M
26	376	AAC GTT	AAT/ ATT N376N/ V377I
27	381	ATC GGT	ATT/ AGT I381I/ G382S
28	398	GCG	GCA/ GTG A398A/ V
29	409	CCG	CCA/ CTG P409P/ L
30	416	CGT	TGT/ CAT R416C/ H
31	421	GCG	GCA/ GTG A421A/ V
32	428	CCG	CCA/ CTG P428P/ L
33	435	TTC GGA	TTT/ AGA F435F/ G436R
34	440*	CGC	TGC/ CAC* R440C/ H
35	449	CGC	TGC/ CAC R449C/ H
36	463	TTC GAC	TTT/ AAC F463F/ D464N
37	486	ATC GAC	ATT/ AAC I486I/ D487N
38	496*	CGC	TGC*/ CAC R496C/ H

Appendix 8- Conserved elements of P450 enzymes

All mitochondrial P450s have a leader peptide of around 20-30 amino acids which is required to target the enzyme to the mitochondria, as discussed in section 1.3.2. In contrast microsomal P450s have a hydrophobic N-terminal signal peptide, which targets them to the membrane of the endoplamic reticulum. The amino acid sequences involved in these N-terminal targetting regions are shown in Figure A8a.

The region of highest amino acid homology in P450 enzyme sequences forms the α-helix, termed helix L, to which the cysteine for the haem ligation is joined (Figure A8b). The cysteinyl thiolate is bracketed on either side by an invariant phenylanine, F350 in CYP101 and a hydrophobic residue, L358 in CYP101, which is either a leucine or isoleucine in microsomal P450s and valine in CYP11A. The close proximity of helix L to the haem disc requires a glycine, the amino acid with the smallest side chain, to be present at the N-terminal of the helix. Six residues in from the invariant glycine, a glutamate is always present. It has been proposed that this conserved glutamate in helix L hydrogen bonds via $3H_2O$ molecules to a threonine residue, which is located a highly conserved amino acid sequence of helix I (Table A8.2) (Poulos *et al*, 1985). This interaction maintains the L and I helices in a parallel orientation, with the haem embedded between the proximal helix L, and distal helix I. This arrangement differs from other haem proteins, such as globins and cytochrome c peroxidase, where the proximal and distal helices are antiparallel and the haem is ligated to a C-terminal histidine (Ten Eyck, 1979; Poulos *et al*, 1980).

The haem propionate and vinyl groups interact with four amino acids, Arg 112, Asp 297, Arg 299 and His 355 in CYP101. The equivalent haem binding residues in the adrenal steroidogenic P450s were obtained from the alignment created for this thesis (see Figure 6.4b) and are shown in Table A8.1 along with those for CYP102. Mutations that alter these key binding residues have been shown to cause loss of enzyme function, such mutations include R448H and R384Q in CYP11B1 and R440H in CYP17 (see section 1.4).

Figure A8a: N-terminal leader peptides of steroidogenic P450s

N-terminal leader peptides of mitochondrial P450s:

```
   +    +    +         +-   + +    -
MLAKGLPPRSVLVKGYQTFLSAPREGLGRLRVPTGEG............CYP11A

   + + -           +      +  +
MALRAKAEVCMAVPWLSLQRAQALGTRAARVPRTVL.............CYP11B1

   + + -           + +    +  +
MALRAKAEVCVAAPWLCLQRARALGTRAARAPRTVL.............CYP11B2
```

N-terminal signal peptides of microsomal P450s:

```
                  +++    +  +
MWELVALLLLTLAYLFWPKRRCPGAKYPKSLLSLP...............CYP17

               +      + +  +
MLLLGLLLLPLLAGARLLWNWWKLRSLHLPPLAPG...............CYP21
```

In the above figure amino acids with basic side chains are marked with a + and those with acidic side chains are marked with a -, to denote charged residues. For single letter amino acid code see Appendix 2.

Figure A8b: Helix L and haem binding region:

```
                         HELIX L

**   *  ** *      *
FGHGSHL-CL-GQHLARREIIVTLKEWLTR...........CYP101    residues 350-377

FGWGVRQ-CL-GRRIAELEMTIFLINMLEN...........CYP11A    residues 455-482

FGFGMRQ-CL-GRRLAEVEMLLLLHHVLKH...........CYP11B1   residues 443-470

FGFGMRQ-CL-GRRLAEVEMLLLLHHVLKR...........CYP11B2   residues 443-470

FGAGPRS-CI-GEILARQELFLIMAWLLQR...........CYP17     residues 435-462

FGCGARV-CL-GEPLARLELFVVLTRLLQA...........CYP21     residues 421-448

FGNGQRA-CI-GQQFALHEATLVLGMMLKH...........CYP102    residues 393-420
```

In the above diagram the helix L region of seven P450s enzymes is shown, with the underlined amino acids indicating the helix. Asterisks * , mark conserved amino acids. The name of the P450 is given to the right, along with the residue numbers for this region.

Table A8.1: Contacts to haem propionate and vinyl groups.

Protein region of contact	P450 Enzyme						
	CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
Helix C	R112	R151	R141	R141	R125	R120	H100
ß-Pair after helix K	D297 R299	L394 R396	L382 R384	L382 R384	I371 H373	L363 H365	L332 K334
Beside thiolate cysteine	H355	R460	R448	R448	R440	R426	R398

Table A8.2: Conserved region of helix I believed to be involved in proton donation to haem.

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
T252	T330	T318	T318	T305	T295	T267
D251	D329	D317	D317	E304	E294	E266
G249	G327	S315	S315	G302	G292	G264
G248	G326	G314	G314	A301	G291	A2478

Table A8.3: Hydrophobic contacts to haem via ß-pair after helix K
(contacts to haem propionate and vinyl group indicated by *).

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
				P367	P360	P328
V295	V392	L380	L380	M368	L361	A329
A296	T393	F381	F881	L369	A362	F330
D297*	L394*	L382*	L382*	I370*	L363*	S331
G298	Q395	E383	E383	P371	P364	L332*
R299*	R396*	R384*	R384*	H372*	H365*	Y333
I300	Y397	V385	V385	K373	R366	A334
L301	L398	A386	V386	A374	T367	K335*

The edge of the haem is surrounded by apolar residues in the interior of the P450 enzyme. Residues interacting with the haem edges extend from one strand of a ß-pair structure, residues 294-299 in CYP101 (Table A8.3). The haem is thus completely surrounded, with no edge accessible to the molecular surface. The closest approach to the haem occurs at the proximal face, a distance of about 8 Å. This corresponds to the upper part of the redox partner docking site. The proximal face of the haem is masked, however, by the hydrophobic residues bracketing the thiolate; equivalent to CYP101 F350 and L358 (Poulos *et al*, 1985). This could suggest that a conformational change occurs on binding of the redox partner that alters the position of these two hydrophobic residues.

A buried haem is also found in other known haem enzyme structures, catalase (Murthy *et al*, 1981) and cytochrome c peroxidase (Poulos *et al*, 1980), but sharply contrasts with the exposed haem edges found in the true cytochromes (Salemme, 1977). This means that unlike cytochromes, P450s must accept electrons via a hydrogen bonding path involving protein groups of the redox partner and P450 enzyme (Poulos *et al*, 1985).

The threonine, T252 of CYP101, which links the two main P450 helices, is usually located four residues away from a glycine, G248, and is always linked to an acidic amino acid (Table A8.2). In microsomal P450s, such as CYP17, the acidic amino acid is glutamic acid, whilst in mitochondrial and most bacterial P450s, including CYP11B enzymes, it is aspartic acid. The aspartic or glutamic acid residue is positioned in close proximity to the haem. It has been suggested that this acidic residue may act to provide one of the protons that catalyze the cleavage of the O-O bond of molecular oxygen (Poulos & Raag, 1992). The other proton required for the reaction may come from the internal solvent channel linking T252 to E366 of CYP101 or other acidic amino acids close to the haem (Raag *et al*, 1991). No internal solvent channel exists in CYP102, yet the threonine and glutamic acid residues are conserved (T268, E267 and E409 in CYP102). It is believed, however, that T268 in CYP102 may donate protons to iron bound oxygen during catalysis. A water molecule lies in a groove in helix I, which is formed by hydrogen bonding

between T268 and the carbonyl oxygens of A264 and G265. This water molecule may accept protons from E267, which is exposed to the solvent in CYP102, pass them on to T268, which then protonates the dioxygen at the haem (Ravichandran *et al*, 1993).

An invariant acidic amino acid residue also occurs at the C-terminal of helix K, E288 in CYP101, three residues away from an invariant arginine, R292 (Table A8.4). This region precedes the ß-pair formed by residues 294-299, which is involved in contacts to the haem. As such the conserved glutamic acid residue would be close to the haem and may be important to the structure of the P450 active site, or possibly play a role in P450 mediated reactions. Another highly conserved region involves a solvent exposed loop above the haem (Table A8.5), which again may be important to the structure of P450 enzymes.

The residues involved in substrate binding define the enzyme specificity and would be expected to vary from enzyme to enzyme. The general features of substrate binding are shown by the recent analysis of CYP102 carried out by Ravichandran *et al* (1993). The active site of CYP102 is accessible through a long hydrophobic channel, lined with hydrophobic residues. This is the substrate binding site. A phenylalanine residue, F87, lies at the bottom of the substrate binding pocket of CYP102, with its phenyl ring almost perpendicular to the porphyrin ring of the haem group. It is thought that this residue may be important in attracting the hydrophobic end of the substrate into the active site. In CYP101 a tyrosine occurs at position 96 within this region and is believed to have the same function of sequestering the substrate to the bottom of the binding pocket (Poulos *et al*, 1985). The adrenal steroidogenic P450s all have a corresponding aromatic or hydrophobic residue in a similar position (Table A8.6).

At the entrance of the substrate binding pocket of CYP102, an exposed hydrophobic patch is formed by the side chains of F11, L14, L17, P18, P45, and A191. These hydrophobic residues may be important for initial docking of a lipophilic substrate since they are solvent-exposed, mobile and located adjacent to the binding pocket. Similar hydrophobic patches may occur in other P450s, but have not been identified to date. Tables A8.6-A8.8 show elements of substrate binding presumed to be involved in substrate binding in the adrenal steroidogenic P450s, CYP101 and CYP102.

Table A8.4: C-terminal of helix K (only invariant residues are numbered).

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
E	K	K	K	R	A	N
E287	E383	E371	E371	E358	E351	E319
L	T	T	T	V	V	A
L290	L385	L373	L373	L361	L353	L322
R291	R386	R374	R374	R362	R354	R323
R	L	L	L	L	L	L
F	H	Y	Y	R	R	W
S	P	P	P	P	P	P

Table A8.5: Solvent exposed loop close to haem (only conserved amino acids are numbered).

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
F	P337	P425	P425	P412	P414	P376
S	T	Q	Q	E	D	E
R342	R439	R427	R427	R416	R408	R378
Q	W	W	W	F	F	F
K	L	L	L	L	L	E
V	S	D	D	N	E	N
S	K	I	I	P	P	P

Table A8.6: Hydrophobic substrate sequestering region before helix C and near haem (amino acids presumed to participate in locating substrate at haem are marked by *).

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
Y96*	G	G	G	G	G	G
D	V	V	V	I	D	L
I	L140*	F130*	F130*	A	Y112*	F87*
F	L	L	L	F114*	S	T
P	K	L	L	A	L	S
T	K	N	N	D	L	W

Table A8.7: Residues involved in substrate specificity on C-terminal end of helix F, which interact by van Der Waals forces with residues of a loop between the C-terminal ß-pair.

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
M	R	W	W	T	G	A
T185	L257	T248	I248	L229	L223	M185

Table A8.8: Residues on helix I, which occur above the haem in the active site.

P450 Enzyme						
CYP101	CYP11A	CYP11B1	CYP11B2	CYP17	CYP21	CYP102
R	A	A	A	T	M	Y
M	N	N	N	T	A	Q
C	V	S	S	I	A	I
G	T	M	M	G	V	I
L	E	E	E	D	D	T
L	M	L	L	I	L	F
L	L	T	T	F	L	L
V	A	A	A	G	I	I
G	G	G	G	A	G	A

The structure of the haemoprotein domain of CYP102 also defines a putative docking site for the NADPH-dependent reductase, a rectangular depression close to the haem. Several conserved arginine and lysine residues that protrude into this depression may be important for docking the reductase, K281, K288, R295 and K308, these correspond to H319, K326, R339 and R348 in CYP17 and H309, R316, H323 and R341 in CYP21. As would be expected these residues are not conserved in CYP101 or the mitochondrial P450s, which use a different redox partner. A second set of lysines and arginines are conserved in these P450s in the same region: corresponding to R344, R353 and K378 in CYP11A.

References

Ahlgren, R., Yanase, T., Simpson, E.R., Winter, J.S.D., & Waterman, M.R. (1992). Compound heterozygous mutations (Arg 239-Stop, Pro 342-Thr) in the CYP17 (P45017α) gene lead to ambiguous external genitalia in a male patient with partial combined 17α-hydroxylase/ 17,20-lyase deficiency. *J. Clin. Endocrinol. Metabol.* **74**:667-672.

Ainsworth, P.J., Surh, L.C., & Coulter-Mackie, M.B. (1991). Diagnostic single strand conformational polymorphism, (SSCP): a simplified non-radioisotopic method as applied to a Tay-Sachs B1 varient. *Nucleic Acids Res.* **19**:405-6.

Amor, M., Parker, K.L., Globerman, H., New, M.I. & White, P.C. (1988). Mutation in the *CYP21B* gene (Ile-Asn) causes steroid 21-hydroxylase defciency. *Proc. Natl. Acad. Sci. USA* **85**:1600-1604.

Azziz, R. (1993). Adrenocortical dysfunction in androgen excess. *Seminars In Reproductive Endocrinology* **2**:353-358.

Bachmann, B., Lüke, W. & Hunsman, G. (1990). Improvement of PCR amplified DNA sequencing with the aid of detergents. *Nucl. Acids Res.* **18**:1309.

Balbas, P., Soberon, X., Merino, E., Zurita, M., Lomeli, H., Valle, F., Flores, N. & Bolivar, F. (1986). Plasmid vector pBR322 and its special-purpose derivatives- A review. *Gene* **50**:3-40.

Baker, K.P. & Schatz, G. (1991). Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria. *Nature* **349**:205-208.

Barnes, R.B., Rosenfield, R.L. & Ehrmann, D.A. (1993). Dysregulation of 17-hydroxylase and 17,20-lyase in functional ovarian hyperandrogenism. *Seminars In Reproductive Endocrinology* **2**:334-340.

Baukal, A.J., Balla, T., Hunyady, L., Hausdorff, W., Guillemette, G. & Catt, K.J. (1988). Angiotensin II and guanine nucleotides stimulate formation of inositol 1,4,5-trisphosphate and its metabolites in permeabilized adrenal glomerulosa cells. *J. Biol. Chem.* **263**:6087-6092.

Becker, A. & Gold, M. (1975). Isolation of the bacteriophage lambda *A*-gene protein. *Proc. Natl. Acad. Sci. USA* **72**:581-585.

Benton, W.D. & Davis, R.W. (1977). Screening lambda gt recombinant clones by hybridization to single plaques *in situ*. *Science* **196**:180-182.

Bérubé, D., Luu-The, V., Lachance, Y., Gagné, R. & Labrie, F. (1989). Assignment of the human 3ß-hydroxysteroid dehydrogenase gene to the p13 band of chromosome 1. *Cytogen. Cell Genet.* **52**:199-200.

Besman, M.J., Yanagibashi, K., Lee, T.D., Kawamura, M., Hall, P.F. & Shively, J.E. (1989). Identification of des-(Gly-Ile)-endozepine as an effector of corticotropin-dependent adrenal steroidogenesis. Stimulation of cholesterol delivery is mediated by the peripheral benzodiazepine receptor. *Proc. Natl. Acad. Sci. USA* **86**:4897-4901.

Bessman, M.J. & Reha-Krantz, L.J. (1977). Studies on the biochemical basis of spontaneous mutation. *J. Mol. Biol.* **116**:115-123.

Biason, A., Mantero, F., Scaroni, C., Simpson, E.R. & Waterman, M.R. (1992). Molecular basis of 17α-hydroxylase/ 17,20 lyase deficiency syndrome. In *Cellular and Molecular Biology of the Adrenal Cortex* (Saez, J.M., Brownie, A.C., Capponi, A., Chambaz, E.M., Mantero, F. eds.) Vol.222, pp.137-145. Colloque INSERM/ John Libbey Eurotext Ltd.

Biglieri, E.G., Herron, M.A. & Brust, N. (1966). 17-Hydroxylation deficiency in man. *J. Clin. Invest.* **45**:1946-1954.

Bird, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucl. Acids Res.* **8**:1499-1504.

Birnboim, H.C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**:1513-1522.

Black, S. (1992). Membrane topology of the mammalian P450 cytochromes. *FASEB J.* **6**:680-685.

Bohr, V.A., Smith, C.A., Okumoto, D.S. & Hanawalt, P.C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO is much more efficient than in the genome overall. *Cell* **40**:359-369.

Bolivar, F., Rodriguez, R.L., Betlach, M.C. & Boyer, H.W. (1977a). Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9. *Gene* **2**:75-93.

Bolivar, F., Rodriguez, R.L., Greene, P.J, Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. & Falkow, S. (1977b). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**:95-113.

Bongiovanni, A.M. & Kellenbenz, G. (1962). The adrenogenital syndrome with deficiency of 3ß-hydroxysteroid dehydrogenase. *J. Clin. Invest.* **41**:2086-2092.

Bongiovanni, A.M. (1981). Acquired adrenal hyperplasia, with special reference to 3ß-HSD. *Fertil. Steril.* **35**:599-608.

Bradshaw, K.D., Waterman, M.R., Couch, R.T., Simpson, E.R & Zuber, M.X. (1987). Characterization of complementary deoxyribonucleic acid for human adrenocortical 17α-hydroxylase: a probe for analysis of 17α-hydroxylase deficiency. *Mol. Endocrinol.* **1**:348-354.

Brown, M.S. & Goldstein, J.L. (1980). Multivalent feedback regulation of HMG CoA reductase, a control mechanism coordinating isoprenoid synthesis and cell growth. *J. Lipid. Res.* **21**:505-517.

Brown, M.S. & Goldstein J.L. (1986). A receptor-mediated pathway for cholesterol homeostasis. *Science* **232**: 34-47.

Cara, J.F., Moshang, T. Jr., Bongiovanni, A.M. & Mark, B.S. (1985). Elevated 17-hydroxyprogesterone and testosterone in a newborn with 3ß-hydroxysteroid dehydrogenase deficiency. *N. Engl. J. Med.* **313**:618-621.

Casey, M.L. & MacDonald, P.C. (1982). Extraadrenal formation of a mineralocorticoid: Deoxycorticosterone and deoxycorticosterone sulfate biosynthesis and metabolism. *Endocr. Rev.* **3**:396-403.

Casey, M.L., Winkel, C.A. & MacDonald, P.C. (1983). Conversion of progesterone to deoxycorticosterone in the human fetus: Steroid 21-hydroxylase activity in fetal tissues. *J. Steroid Biochem.* **18**:449-452.

Chaubert, P., Dolores, B. & Benhattar, J. (1993). An improved method for rapid screening of DNA mutations by nonradioactive Single-Strand Conformation Polymorphism Procedure. *Biotechniques* **15**:116.

Chiou, S.H., Hu, M.C. & Chung, B.C. (1990). A missense mutation at $Ile^{172} \rightarrow Asn$ or $Arg^{356} \rightarrow Trp$ causes steroid 21-hydroxylase deficiency. *J. Biol.Chem.* **265**:3549-3552.

Chua, S.C., Szabo, P., Vitek, A., Grzeschik, K.-H., John, M., & White, P.C. (1987). Cloning of cDNA encoding steroid 11ß-hydroxylase ($P450_{c11}$). *Proc. Natl. Acad. Sci. USA* **84**:7193-7197.

Chung, B., Matteson, K.J., Voutilainen, R., Mohandas, T.K. & Miller, W.L. (1986). Human cholesterol side chain cleavage enzyme, P450scc: cDNA cloning, assignment of the gene to chromosome 15, and expression in the placenta. *Proc. Natl. Acad. Sci. USA* **83**:8962-8966.

Chung, B., Picado-Leonard, J., Haniu, M., Bienkowski, M., Hall, P.F., Shivley, J.E. & Miller, W.L. (1987). Cytochrome P450c17 (steroid 17α-hydroxylase/17,20 lyase): Cloning of human adrenal and testis cDNAs indicate the same gene is expressed in both tissues. *Proc. Natl. Acad. Sci. USA* **84**:407-411.

Cicila, G.T., Rapp, J.P., Wang, J.-M., St. Lezin, E., Ng, S.C. & Kurtz, T.W. (1993). Linkage of 11ß-hydroxylase mutations with altered steroid biosynthesis and blood pressure in the Dahl rat. *Nature Genet.* **3**:346-353.

Clark, A.J., Balla, T., Jones, M.R. & Catt, K.J. (1992). Stimulation of early gene expression by angiotensin II in bovine adrenal glomerulosa cells: roles of calcium and protein kinase C. *Mol. Endocrinol.* **6**:1889-1898.

Coghlan, V.M. & Vickery, L.E. (1991). Site-specific mutations in human ferredoxin that affect binding to ferredoxin reductase and cytochrome $P450_{scc}$. *J. Biol. Chem.* **266**:18606-18612.

Cohen, S.N., & Chang, A.C.Y. (1972). Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci.* **69**:2110-2114.

Cohen, S.N, Chang, A.C.Y., Boyer, H.W. & Helling, R.B. (1973). Construction of biologically functional bacterial plasmids *in vitro*. *Proc. Natl. Acad. Sci. USA* **70**:3240-3244.

Coon, M.J. & White, R.E. (1980). In *Dioxygen binding and activation by metal centers* (Spiro, T.G. ed.) pp.73-123. Wiley, New York.

Cooper, D.N. & Krawczak, M. (1993). Single base substitutions. In *Human Gene Mutation* pp.109-153. BIOS Scientific Publishers Ltd., Oxford.

Cotton, R.G.H., Rodriguez, N.R. & Campbell, R.D. (1988). Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc. Natl. Acad. Sci. USA* **85**:4397-401.

Covey, C., Richardson, D. & Carbon, J. (1976). A method for the deletion of restriction sites in bacterial plasmid deoxyribonucleic acid. *Mol. Gen. Genet.* **145**:155-158.

Cupp, J.R. & Vickery, L.E. (1988). Identification of free and $[Fe_2\text{-}S_2]$-bound cysteine residues of adrenodoxin. *J. Biol. Chem.* **263**:17418-17421.

Curnow, K.M., Tusie-Luna, M.T., Pascoe, L., Natarajan, R., Gu, J.L., Nadler, J.L. & White, P.C. (1991). The product of the *CYP11B2* gene is required for aldosterone biosynthesis in the human adrenal cortex. *Mol. Endocrinol.* **5**:1513-1521.

Curnow, K.M., Slutsker, L., Vitek, J., Cole, T., Speiser, P.W., New, M.I., White, P.C. & Pascoe, L. (1993). Mutations in the *CYP11B1* gene causing congenital adrenal hyperplasia and hypertension cluster in exons 6, 7 and 8. *Proc. Natl. Acad. Sci. USA* **90**:4552-4556.

Demers, D.B., Odelberg, S.J., & Fisher, L.M. (1990). Identification of a factor IX point mutation using SSCP analysis and direct sequencing. *Nucl. Acids Res.* **18**:5575.

Deshaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A. & Schekman, R. (1988). A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. *Nature* **332**:800-805.

de Peretti, E. & Forest, M.G. (1982). Pitfalls in the etiological diagnosis of congenital adrenal hyperplasia in the early neonatal period. *Horm. Res.* **16**:10-22.

Di Blasio, A.M., Voutilainen, R.,Jaffe, R.B., & Miller, W.L. (1987). Hormonal regulation of mRNA's for P450scc (cholesterol side-chain cleavage enzyme) and P450c17 (17α-hydroxylase/17,20 lyase) in cultured human fetal adrenal cells. *J. Clin. Endocrinol. Metab.* **65**:170-175.

Dignam, J.D. & Strobel, H.W. (1977). NADPH-cytochrome P-450 reductase from rat liver, purification by affinity chromatography and characterization. *Biochemistry* **16**:1116-1122.

Dockhorn-Dworniczak, B., Dworniczak, B., Brömmelkamp, L., Bülles, J., Horst, J. & Böcker, W.W. (1991). Non-isotopic detection of single strand conformation polymorphism (PCR-SSCP): a rapid and sensitive technique in diagnosis of phenylketonuria. *Nucl. Acids Res.* **19**:2500.

Domalik, L.J., Chaplin, D.D., Kirkman, M.S., Wu, R.C., Liu, W., Howard, T.A., Seldin, M.F. & Parker, K.L. (1991). Different isozymes of mouse 11ß-hydroxylase produce mineralocorticoids and glucocorticoids. *Mol. Endocrinol.* **5**:1853-1861.

Ehrlich, H.A (1989). Protocols. In *PCR Technology: Principles and Applications* (Ehrlich, H.A. ed.) pp.35-36. Stockton Press, New York, USA.

Ehrlich, M. & Wang, R.Y.-H. (1981). 5-Methylcytosine in eukaryotic DNA. *Science* **212**:1350-1357.

Embley, T.M. (1991). The linear PCR reaction: a simple and robust method for sequencing amplified rRNA genes. *Lett. Applied Microbiol.* **13**:171-174.

Enyedi, P., Büki, B., Mucsi, I. & Spät, A. (1985). Polyphosphoinositide metabolism in adrenal glomerulosa cells. *Mol. Cell. Endocrinol.* **41**:105-112.

Enyedi, P., Mucsi, I., Hunyady, L., Catt, K.J. & Spät, A. (1986). The role of guanyl nucleotide binding proteins in the formation of inositol phosphates in adrenal glomerulosa cells. *Biochem. Biophys. Res. Commun.* **140**:941-947.

Estabrook, R.W., Cooper, D.Y. & Rosenthal, O. (1963). The light reversible carbon monoxide inhibition of the steroid C21-hydroxylase system of the adrenal cortex. *Biochem. Z.* **338**:741-755.

Fardella, C.E., Zhang, L.-H., Mahachoklertwattana, P., Lin, D. & Miller, W.L. (1993). Deletion of amino acids Asp^{487}-Ser^{488}-Phe^{489} in human cytochrome P450c17 causes severe 17α-hydroxylase deficiency. *J. Clin. Endocrinol. Metab.* **77**:489-493.

Fardella, C.E., Hum, D.W., Homoki, J. & Miller, W.L. (1994a). Point mutation of Arg440 to His in cytochrome P450c17 causes severe 17α-hydroxylase deficiency. *J. Clin. Endocrinol. Metab.* **79**:160-164.

Fardella, C.E., Hum, D.W & Miller, W.L. (1994b). Gene conversion in exons 3 and 4 of the *CYP11B2* gene encoding P450c11AS causes primary hypoaldosteronism (CMO II) deficiency. Abstract 218, 76th. Annual Meeting of The Endocrine Society, June 1994.

Fischer, S.G. & Lerman, L.S. (1980). Separation of random fragments of DNA according to properties of their sequences. *Proc. Natl. Acad. Sci. USA* **77**:4420-4424

Fraser, R., Brown, J.J., Mason, P.A., Morton, J.J., Lever, A.F., Robertson, J.I.S., Lee, H.A. & Miller, H. (1987). Severe hypertension with absent secondary sex characteristics due to deficiency of steroid 17α-hydroxylase activity. *J. Hum. Hypertension* **1**:53-58.

Freeman, W.H. & Bracegirdle, B. (1985). In *An advanced atlas of histology*, p.139. Heineman Educational Books, London.

Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. (1983). Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* **170**:827-842.

Garfinkel, D. (1958). Studies on pig microsomes. I. Enzymatic and pigment composition of different microsomal fractions. *Arch. Biochem. Biophys.* **77**:493-509.

Geren, L.M., O'Brien, P., Stoneheuerner, J. & Millet, F. (1984). Identification of specific carboxylate groups on adrenodoxin that are involved in the interaction with adrenodoxin reductase. *J. Biol. Chem.* **259**:2155-2160.

Gillette, J.R. (1966). Biochemistry of drug oxidants and reduction by enzymes in hepatic endoplasmic reticulum. *Adv. Pharmacol.* **4**:219-261.

Globerman, H., Amor, M., Parker, K.L., New, M.I. & White, P.C. (1988). Nonsense mutation causing steroid 21-hydroxylase deficiency. *J.Clin.Invest.* **82**:139-144.

Gnanaiah, W. & Omdahl, J.L. (1986). Isolation and characterization of pig kidney mitochondrial ferredoxin:$NADP^+$ oxidoreductase. *J. Biol. Chem.* **261**:12649-12654.

Goebelsmann, U., Zachmann, M. & Davajan, V. (1976). Male pseudohermaphroditism consistent with 17,20 desmolase deficiency. *Gynecol. Invest.* **7**:138-156.

Goldsmith, O., Solomon, D.H. & Horton, R. (1976). Hypogonadism and mineralocorticoid excess. The 17-hydroxylase deficiency syndrome. *N. Engl. J. Med.* **277**:673-677.

Golos, T.G. & Strauss J.F. (1988). III 8-Bromo adenosine cyclic 3'5'phosphate rapidly increases 3-hydroxy-3-methylglutaryl coenzyme A reductase mRNA in human granulosa cells: role of cellular sterol balance in controlling the response to tropic stimulation. *Biochemistry* **27**:3503-3506.

Gonzalez, G.A. & Montiminy, M.R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* **59**:675-680.

Grunstein, M., & Hogness, D.S. (1975). Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA* **72**:3961-3965.

Gunsalus, I.C., Tyson, C.A., & Lipscomb, J.D. (1973). In *Oxidases and Related Redox Systems* (King, T.E., Mason, H.S. & Morrison, M., eds.) Vol. 2. p583. University Park Press, Baltimore, USA.

Hahn, R.A., Teutsch, S.M., Rothenberg, R.B. & Marks, J.S. (1990). Excess deaths from nine chronic diseases in the United States, 1986. *J. Am. Med. Assoc.* **264**:2654-2659.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557-580.

Handler, J.D., Schimmer, B.P., Flynn, T.R., Szyf, M., Seidman, J.G. & Parker, K.L. (1988). An enhancer element and a functional cyclic AMP-dependent protein kinase are required for expression of adrenocortical 21-hydroxylase. *J. Biol. Chem.* **263**:13068-13073.

Harada, F., Kimura, A., Iwanaga, T., Shimozawa, K., Yata, J. & Sasazuki, T. (1987). Gene conversion-like events cause steroid 21-hydroxylase deficiency in congenital adrenal hyperplasia. *Proc. Natl. Acad. Sci. USA* **84**:8091-8094.

Hartl, F.U. & Neupert, W. (1990). Protein sorting to mitochondria: evolutionary conservations of folding and assembly. *Science* **247**:930-938.

Hauffa, B.P., Miller, W.L., Grumbach, M.M., Conte, F.A. & Kaplan, S.L. (1985). Congenital adrenal hyperplasia due to deficient side-chain cleavage activity (20,22 desmolase) in a patient treated for 18 years. *Clin. Endocrinol.* **23**:481-493.

Hayashi, K. (1992). PCR-SSCP: A simple and sensitive method for detection of mutations in genomic DNA. *PCR Methods Appl.* **1**:34-38.

Helmberg, A., Ausserer, B. & Kofler, R. (1992). Frame shift by insertion of two base pairs in codon 394 of *CYP11B1* causes congenital adrenal hyperplasia due to steroid 11ß-hydroxylase deficiency. *J. Clin. Endocrinol. Metab.* **75**:1278-1281.

Hershfield, V., Boyer, H.W., Yanofsky, C., Lovett, M.A. & Helinski, D.R. (1974). Plasmid ColE1 as a molecular vehicle for cloning and amplification of DNA. *Proc. Natl. Acad. Sci. USA* **71**:3455-3459.

Higahsi, Y., Yoshioka, H., Yamana, M., Gotoh, O. & Fujii-Kurijama, Y. (1986). Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome 6: a pseudogene and a genuine gene. *Proc. Natl. Acad. Sci. USA* **83**:2841-2845.

Higashi, Y., Tanae, A., Inoue, H., Hiromasa, T. & Fujii-Kurijama, Y. (1988). Aberrant splicing and missense mutation causing steroid 21-hydroxylase (P-450c21) deficiency in humans: possible gene conversion products. *Proc. Natl. Acad. Sci. USA* **85**:7486-7490.

Higashi, Y., Hiromasa, T., Tanae, A., Miki, T., Nakura, J., Kondo, T., Ohura, T., Ogawa, E., Nakayama, K. & Fujii-Kuriyama, Y. (1991). Effects of individual mutations in the P-450(C21) pseudogene on the P-450(C21) activity and their distribution in the patient genomes of congenital steroid 21-hydroxylase deficiency. *J. Biochem.* **109**:638-644.

Hogg, A., Onadim, Z., Baird, P.N. & Cowell, J.K. (1992). Detection of heterozygous mutations in the *RB1* gene in retinoblastoma patients using single-strand conformation polymorphism analysis and polymerase chain reaction sequencing. *Oncogene* **7**:1445-1451.

Honda, S.I., Morohashi, K.I. & Omura, T. (1990). Novel cAMP regulatory elements in the promoter region of bovine P-450 (11ß) gene. *J. Biochem.* (Tokyo) **108**:1042-1049.

Horwitz, J.P., Chua, J., Curby, R.J., Tomson, A.J., DaRooge, M.A., Fisher, B.E., Mauricio, J. & Klundt, I. (1964). Substrates for cytochemical demonstration of enzyme activity. I. Some substituted 3-indoyl-ß-D-glycopyranosides. *J. Med. Chem.* **7**:574-575.

Hsueh, A.J., Peck, E.J. & Clark, J.H. (1976). Control of uterine estrogen receptor levels by progesterone. *Endocrinology* **98**:438-444.

Huang, C.H., Reid, M., Daniels, G. & Blumenfeld, O.O. (1993). Alteration of splice site selection by an exon mutation in the human glycophorin A gene. *J. Biol. Chem.* **268**:25902-25908.

Huynh, T.V., Young, R.A. & Davis, R.W. (1985). Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. In *DNA cloning: A practical approach* (ed. D.M. Glover), vol. 1, p.49. IRL Press, Oxford.

Hwang, S.T. & Schatz, G. (1989). Translocation of proteins across the mitochondrial inner membrane, but not into the outer membrane, requires nucleoside triphosphates in the matrix. *Proc. Natl. Acad. Sci. USA* **86**:8432-8436.

Imai, M., Shimada, H. Okada, Y, Matsushima-Hibiya, Y., Ogishima, T. & Ishimura, Y. (1990). Molecular cloning of a cDNA encoding aldosterone synthase cytochrome P-450 in rat adrenal cortex. *FEBS Lett.* **263**:299-302.

Imai, T., Globerman, H., Gertner, J.M., Kagawa, N. & Waterman, M.R. (1993). Expression and purification of functional human 17α-hydroxylase/17,20-lyase (P450c17) in *Escherichia coli*. *J. Biol. Chem.* **268**:19681-19689.

Ish-Horowicz, D. & Burke, J.F. (1981). Rapid and efficient cosmid cloning. *Nucl. Acids Res.* **9**:2989-2998.

Iyanagi, T., Makino, N. & Mason, H.S. (1974). Redox properties of the reduced nicotinamide adenine dinucleotide phosphate-cytochrome P-450 and reduced nicotinamide adenine dinucleotide-cytochrome b_5 reductases. *Biochemistry* **13**:1701-1710.

Jeanpierre, M. (1989). A rapid method for the purification of DNA from blood. *Nucl. Acids Res.* **15**:9611.

Jefcoate, C.R., McNamara, B.C. & DiBartolomeis, M.S. (1986). Control of steroid synthesis in adrenal fasciculata. *Endocr. Res.* **12**:315-350

John, M.E., John, M.C., Boggaram, V., Simpson, E.R. & Waterman, M.R. (1986). Transcriptional regulation of steroid hydroxylase genes by ACTH. *Proc. Natl. Acad. Sci. USA* **83**:4715-4719.

Kagimoto, M., Winter, J.S.D., Kagimoto, K., Simpson, E.R. & Waterman, M.R. (1988). Structural characterization of normal and mutant steroid 17α-hydroxylase genes: molecular basis of one example of combined 17α-hydroxylase/ 17,20 lyase deficiency. *Mol. Endocrinol.* **2**:564-570.

Kaiser, K. & Murray, N.E. (1985). The use of phage lambda replacement vectors in the construction of representative genomic DNA librairies. In DNA cloning: A practical approach (Glover, D.M., ed.), vol. 1, IRL Press, Oxford.

Kang, P.-J., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A. & Pfanner, N. (1990). Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. *Nature* **348**:137-143.

Kato, R., Iwasaki, K., Shiraga, T. & Noguchi, H. (1976a). Evidence for the involvement of cytochrome P-450 in reduction of benzo(a)pyrene 4,5-oxide by rat liver. *Biochem. Biophys. Res. Commun.* **70**:681-687.

Kato, R., Iwasaki, K. & Noguchi, H. (1976b). Stimulatory effect of FMN and metyl viologen on cytochrome P-450 dependent reduction of tertiary amine N-oxide. *Biochem. Biophys. Res. Commun.* **72**:267-274.

Kawahara, Y., Sunako, M., Tsuda, T., Fukuzaki, H., Fukumoto, Y. & Takai, Y. (1988). Angiotensin II induces expression of the c-fos gene through protein kinase C activation and calcium mobilization in cultured vascular smooth muscle cells. *Biochem. Biophys. Res. Commun.* **150**:52-59.

Kawamoto, T., Mitsuuchi, Y., Toda, K., Miyahara, K., Yokoyama, Y., Nakao, K., Hosoda, K., Yomamoto, Y., Imura, H. & Shizuta, Y. (1990). Cloning of cDNA and genomic DNA for human cytochrome P-$450_{11ß}$. *FEBS Lett.* **269**:345-349.

Kawamoto, T., Mitsuuchi, Y., Toda, Y., Yokoyama, Y., Miyahara, K., Miura, S., Onishi, T., Ichikawa, Y., Nakao, K., Imura, H., Ulick, S. & Shizuta, Y. (1992). Role of steroid 11ß-hydroxylase and steroid 18-hydroxylase in the biosynthesis of glucocorticoids and mineralocorticoids in humans. *Proc. Natl. Acad. Sci. USA* **89**:1458-1462.

Keen, J., Lester, D., Inglehearn, C., Curtis, A. & Bhattacharya, S. (1991). Rapid detection of single base mismatches as heteroduplexes on Hydrolink gels. *Trends Genet.* **7**:5.

Keeney, D.S. & Waterman, M.R. (1993). Regulation of steroid hydroxylase gene expression: importance to physiology and disease. *Pharmac. Ther.* **58**:301-317.

Keohavong, P. & Thilly, W.G. (1989). Fidelity of DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **86**:9253-9257.

Klingenberg, M. (1958). Pigments of rat liver microsomes. *Arch. Biochem. Biophys.* **75**:376-386.

Koh, Y., Buczko, E. & Dufau, M.L. (1993). Requirement of phenylalanine 343 for the preferential Δ^4-lyase versus Δ^5-lyase activity of rat CYP17. *J. Biol. Chem.* **268**:18267-18271.

Kojima, I., Kojima, K., Kreutter, D. & Rasmussen, H. (1984). The temporal integration of the aldosterone secretory response to angiotensin occurs via two intracellular pathways. *J. Biol. Chem.* **259**:14448-14457.

Kominami, S., Ochi, H., Kobayashi, Y., & Takemori, S. (1980). Studies on the steroid hydroxylation system in adrenal cortex microsomes: Purification and characterization of cytochrome P450 specific for steroid 21-hydroxylation. *J. Biol. Chem.* **255**:3386-3394.

Labrie, F., Simard, J., Luu-The, V., Pelletier, G. & Bélanger, A. (1992). Cloning, expression and regulation of tissue-specific expression of 3ß-hydroxysteroid dehydrogenase/ Δ^5-Δ^4 isomerase. In *Cellular and Molecular Biology of the Adrenal Cortex* (Saez, J.M., Brownie, A.C., Capponi, A., Chambaz, E.M., Mantero, F. eds.) Vol, 222 pp. 89-109. Colloque INSERM/ John Libbey Eurotext Ltd.

Lachance, Y., Luu-The, V., Verrault, H., Dumont, M., Rhéaume, E., Lablanc, G. & Labrie, F. (1991). Structure of human type II 3ß-hydroxysteroid dehydrogenase/ Δ^5-Δ^4 isomerase (3ß-HSD) gene: adrenal and gonadal specificity. *DNA Cell Biol.* **10**:701-711.

Lambeth, J.D., Seybert, D. & Kamin, H (1979). Ionic effects of adrenal steroidogenic electron transport: The role of adrenodoxin as an electron shuttle. *J. Biol. Chem.* **254**:7255-7264.

Lambeth, J.D. & Pember, S.O. (1983). Cytochrome P-450scc- adrenodoxin complex: reduction properties of the substrate associated cytochrome and relation of the reduction states of heme and iron-sulfur centers to association of the proteins. *J. Biol. Chem.* **258**:5596-5602.

Lauber, M. & Müller, J. (1989). Purification and characterization of two distinct forms of rat adrenal cytochrome $P450_{11ß}$: functional and structural aspects. *Arch. Biochem. Biophys.* **274**:109-119.

Laughton, C.A., Neidle, S., Zvelebil, M.J.J.M. & Sternberg, M.J.E. (1990). A molecular model for the enzyme cytochrome $P450_{17\alpha}$, a major target for the chemotherapy of prostatic cancer. *Biochem. Biophys. Res. Commun.* **171**:1160-1167.

Lench, N., Stanier, P. & Williamson, R. (1988). Simple non-invasive method to obtain DNA for gene analysis. *The Lancet* June 18, 1356-1358.

Lewin, B. (1990). In *Genes IV.* Mechanisms of splicing, pp. 598-600. Oxford University Press, Oxford.

Levinson, G. & Gutman, G.A. (1987). Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**:203-221

Lifton, R.P., Dluhy, R.G., Powers, M., Rich, G.M., Cook, S., Ulick, S. & Lalouel, J.-M. (1992). Hereditary hypertension caused by chimaeric gene duplications and ectopic expression of aldosterone synthase. *Nature Genet.* **2**:66-74.

Lifton, R.P. & Dluhy, R.G. (1993). The molecular basis of a hereditary form of hypertension, glucocorticoid-remediable aldosteronism. *Trends in Endocrinology and Metabolism* **4**:57-61.

Lin, D., Harikrishna, J.A., Moore, C.C.D., Jones, K.L. & Miller, W.L. (1991a). Missense mutation $Serine^{106} \rightarrow$ Proline causes 17α-hydroxylase deficiency. *J. Biol. Chem.* **266**:15992-15998.

Lin, D., Gitelman, S.E., Saenger, P., Jones, K.L. & Miller, W.L. (1991b). Normal genes for the cholesterol side chain cleavage enzyme, P450scc, in congenital lipoid adrenal hyperplasia. *J. Clin. Invest.* **88**:1955-1962.

Lin, D., Zhang, L.-H., Chiao, E. & Miller, W.L. (1993). Modeling and mutagenesis of the active site of human P450c17. *Mol. Endocrinol.* **8**:392-402.

Lindberg, R. & Negishi, M. (1989). Alteration of mouse cytochrome $P450_{coh}$ substrate specificity by mutation of a sinle amino-acid residue. *Nature* **339**:632-634.

Lucky, A.W., Rosenfield, R.L., McGuire, J., Rudy, S. & Helke, J. (1986). Adrenal androgen hyperresponsiveness to adrenocorticotropin in women with acne and/ or hirsutism: Adrenal enzyme defects and exaggerated adrenarche. *J. Clin. Endocrinol. Metab.* **62**:840-848.

Mandel, M. & Higa, A. (1970). Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**:159-162.

Maniatis, T. & Reed, R. (1987). The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325**:673-678.

Mashiyama, S., Sekiya, T. & Hayashi, K. (1990). Screening of multiple DNA samples for detection of sequence changes. *Techniques* **2**:304-306.

Matsukawa, N., Nonaka, Y., Ying, Z., Higaki, J., Ogihara, T. & Okamoto, M. (1990). Molecular cloning and expression of cDNAs encoding rat aldosterone synthase: variants of cytochrome P-45011ß. *Biochem. Biophys. Res. Commun.* **169**:245-252.

Matteson, K.J., Picado-Leonard, J., Chung, B.-C., Mohandas, T.K. & Miller, W.L. (1986a). Assignment of the gene for adrenal P450c17 (17α-hydroxylase/17,20 lyase) to human chromosome 10. *J. Clin. Endocrinol. Metab.* **63**:789-791.

Matteson, K.J., Chung, B.-C., Urdea, M.S. & Miller, W.L. (1986b). Study of cholesterol side chain cleavage (20,22 desmolase) deficiency causing congenital lipoid adrenal hyperplasia, using bovine P450scc oligodeoxynucleotide probes. *Endocrinology* **118**:1296-1305.

Meadus, W.J., Mason, J.I. & Squires, E.J. (1993). Cytochrome P450c17 from porcine and bovine adrenal catalyses the formation of 5,16-andostadien-3ß-ol from pregnenolone in the presence of cytochrome b_5. *J. Steroid. Biochem. Molec. Biol.* **46**:565-572.

Mellon, S.H. & Miller, W.L. (1989). Extra-adrenal steroid 21-hydroxylation is not mediated by P450c21. *J. Clin. Invest.* **84**:1497-1502.

Mellon, S.H., Shively, J.E. & Miller, W.L. (1991). Human proopiomelanocortin (79-86), a proposed androgen stimulatory hormone, does not affect steroidogenesis in cultured human fetal adrenal cells. *J. Clin. Endocrinol. Metab.* **72**:19-22.

Merrill, C.R., Goldman, D., Sedman, S.A., & Ebert, M.H. (1981). Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. *Science* **211**:1437-1438.

Miller, W.L. (1991). Congenital Adrenal Hyperplasia. *Endocrinol. Metab. Clin. North. Am.* **20**:721-749.

Miura, S. & Ichikawa, Y. (1991a). Proton nuclear magnetic resonance investigation of adrenodoxin. *Eur. J. Biochem.* **197**:747-757.

Miura, S. & Ichigawa, Y. (1991b). Conformational change of adrenodoxin induced by reduction of iron-sulphur cluster. *J. Biol. Chem.* **266**:6252-6258.

Miura, S., Tomita, S. & Ichikawa, Y. (1991). Modification of histidine 56 in adrenodoxin with pyrocarbonate inhibited the interaction with cytochrome P450scc and adrenodoxin reductase. *J. Biol. Chem.* **256**:19212-19216.

Mohabeer, A.J., Hiti, A.L., and Martin, W.J. (1991). Non-radioactive single strand conformation polymorphism (SSCP) using the Pharmacia 'PhastSystem'. *Nucl. Acids Res.* **19**:3154.

Monier, S. van Luc, P., Kreibich, G., Sabatini, D.D. & Adesnik, M. (1988). Signals for the incorporation and orientation of cytochrome P-450 in the endoplasmic reticulum membrane. *J. Cell. Biol.* **107**:457-470.

Monno, S., Ogawa, H., Date, T., Fujioka, M., Miller, W.L. & Kobayashi, M. (1993). Mutation of histidine 373 to leucine in cytochrome P450c17 causes 17α-hydroxylase deficiency. *J. Biol. Chem.* **268**:25811-25817.

Moore, C.C.D. & Miller, W.L. (1991). The role of transcriptional regulation in steroid hormone biosynthesis. *J. Steroid. Biochem.* **40**:517-525.

Morel, Y., Bristow, J., Gitlelman, S.E. & Miller, W.L. (1989). Transcript encoded on the opposite strand of human steroid 21-hydroxylase/ complement component C4 gene locus. *Proc. Natl. Acad. Sci. USA* **86**:6582-6586.

Morel, Y. & Miller, W.L. (1991). Clinical and molecular genetics of congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Adv. Hum. Genet.* **20**:1-68.

Mornet, E., Dupont, J., Vitek, A., & White, P.C. (1989). Characterization of two genes encoding human steroid 11ß-hydroxylase ($P450_{11ß}$). *J. Biol. Chem.* **264**:20961-20967.

Mornet, E., Crété, P., Kuttenn, F., Raux-Demay, M.C., Boué, J., White, P.C. & Boué, A. (1991). Distribution of deletions and seven point mutations on *CYP21B* genes in three clinical forms of steroid 21-hydroxylase deficiency. *J. Hum. Genet.* **48**:79-88.

Morohashi, K., Sogawa, K., Omura, T. & Fujii-Kuriyama, Y. (1987). Gene structure of human cytochrome P450(SCC), cholesterol desmolase. *J. Biochem.* **101**:879-887.

Morohashi, K., Nonaka, Y., Kirita, S., Hatano, O., Takakusu, A., Okamoto, M. & Omura, T. (1990). Enzymatic activities of P-450(11ß)s expressed by two cDNAs in COS-7 cells. *J. Biochem.* **107**:635-640.

Morohashi, K., Honda, S., Inomata, Y., Handa, H. & Omura, T. (1992). A common *trans*-acting factor, Ad4-binding protein, to the promoters of steroidogenic P450s. *J. Biol. Chem.* **267**:17913-17919.

Murakami, Y., Katahira, M., Makino, R., Hayachi, K., Hirohashi, S. & Sekiya, T. (1991). Inactivation of the retinoblastoma gene in a human lung carcinoma cell line detected by single-strand conformation polymorphism analysis of the polymerase chain reaction product of cDNA. *Oncogene* **6**:37-42.

Murthy, M.R.N., Reid, T.J., Sicignano, A., Tanaka, N. & Rossman, M.G. (1981). Structure of beef liver catalase. *J. Mol. Biol.* **152**:465-499.

Myers, R.M., Larin, Z. & Maniatis, T. (1985). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA: DNA duplexes. *Science* **230**:1242-1246.

Myers, R.M., Maniatis, T. & Lerman, L.S. (1987). Detection and localization of single base changes by denaturing gel electrophoresis. *Methods Enzymol.* **155**:501-527.

Naiki, Y., Kawamoto, T., Mitsuuchi, Y., Miyahara, K., Toda, K., Orii, T., Imura, H. & Shizuta, Y. (1993). A nonsense mutation (TGG [Trp^{116}] → TAG [Stop]) in *CYP11B1* causes steroid 11ß-hydroyxlase deficiency. *J. Clin. Endocrinol. Metab.* **77**:1677-1682.

Nakajin, S., Shively, J.E., Yuan, P.M. & Hall, P.F. (1981). Microsomal P450 from neonatal pig testis: Two enzymatic activities (17α-hydroxylase and 17,20 lyase) associated with one protein. *Biochemistry* **20**:4037-4042.

Nebert, D.W., Jones, J.E., Owens, J. & Puga, A. (1988). In *Oxidases and Related Redox Systems* (King, T.E., Mason, H.S. & Morrison, M. eds.) pp.557-576. Liss, New York, USA.

Nelson, D.R. & Strobel, H.W. (1987). Evolution of Cytochrome P-450 Proteins. *Mol. Biol. Evol.* **4**:572-593.

Nelson, D.R., Kamataki, T., Waxman, D.J., Guengerich, F.P., Estabrook, R.W., Feyereisen, R., Gonzalez, F.J., Coon, M.J., Gunsalus, I.C., Gotoh, O., Okuda, K. & Nebert, D.W. (1993). The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol.* **12**:1-51.

New, M.I., Dupont, B., Grumbach, K. & Levin, L.S. (1982). Congenital adrenal hyperplasia and related conditions. In *The Metabolic Basis of Inherited Disease*, 5th. ed. (Stanbury, J.B.S., Wyngarden, J.B., Frederickson, D.S., Goldstein, J.L. & Brown, M.S. eds.) pp973-1000. McGraw-Hill, New York, USA.

New, M.I., White, P.C., Pang, S., Dupont, B. & Speiser, P.W. (1989). The adrenal hyperplasias. In *The metabolic basis of inherited disease*, 6th. ed. (Scriver, B., Beaudet, A.L., Sly, W.S. & Valle, D. eds.) pp.1881-1897. McGraw-Hill, New York, USA.

Norrander, J., Kempe, T. & Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**:101-106.

Ogishima, T., Mitani, F. & Ishimura, Y. (1989). Isolation of two distinct cytochromes P-45011ß with aldosterone synthase activity from bovine adrenocortical mitochondria. *J. Biochem.* (Tokyo) **105**:497-499.

Ogishima, T., Shibata, H., Shimada, H., Mitani, F., Suzuki, H., Saruta, T. & Ishimura, T. (1991). Aldosterone synthase cytochrome P-450 expressedin the adrenals of patients with primary aldosteronism. *J. Biol. Chem.* **266**:10731-10734.

Ohnishi, T., Wada, A., Lauber, M., Yamano, T. & Okamoto, M. (1988). Aldosterone biosynthesis in mitochondria of isolated zones of adrenal cortex. *J. Steroid. Biochem.* **31**:73-81.

Omura, T. & Sato, R. (1964). The carbon monoxide-binding pigment of liver microsomes. *J. Biol. Chem.* **239**:2370-2378.

Onoda, M. & Hall, P.F. (1982). Cytochrome b_5 stimulates purified testicular microsomal cytochrome P450 (C21 side-chain cleavage). *Biochem. Biophys. Res. Commun.* **108**:454.

O'Riordan, J.L.H., Malan, P.G. & Gould, R.P. (1985). In *Essentials of Endocrinology*, pp.78-80. Blackwell Scientific Publications, Oxford.

Orita, M., Iwahara, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989a). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA* **86**:2766-2770.

Orita, M.Y., Suzuki, Y., Sekiya, T. & Hayashi, K., (1989b). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**:874-879.

Orita, M., Sekiya, T. & Hayashi, K. (1990). DNA sequence polymorphisms in *Alu* repeats.*Genomics* **8**:271-278.

Osterman, J., Horwich, A.L., Neupert, W. & Hartl, F.-U. (1989). Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. *Nature* **341**:125-130.

Pang, S., Levine, L.S., Stoner, E., Opitz, J.M., Pollack, M.S., Dupont, B. & New, M.I. (1983). Nonsalt-losing congenital adrenal hyperplasia due to 3ß-hydroxysteroid dehydrogenase deficiency with normal glomerulosa function. *J. Clin. Endocrinol. Metab.* **56**:808-819.

Parissenti, A., Parker, K.L. & Schimmer, B.P. (1993). Identification of promoter elements in the mouse 21-hydroxylase (Cyp21) gene that require a functional cAMP-dependent protein kinase. *Mol. Endocrinol.* **7**:283-290.

Parker, L.N. & Odell, W.D. (1980). Control of adrenal androgen secretion. *Endocr. Rev.* **1**:392-410.

Parker, L.N., Litrak, E.T. & Odell, W.D. (1983). A 60,000 molecular weight human pituitary glycopeptide stimulates adrenal androgen secretion. *Endocrinology* **113**:2092-2096.

Parker, K.L. & Schimmer, B.P. (1993). Transcriptional regulation of the adrenal steroidogenic enzymes. *Trends in Endocrinology and Metabolism* **4**:46-50.

Pascoe, L., Curnow, K.M., Slutsker, L., Connell, J.M.C., Speiser, P., New, M.I. & White, P.C. (1992a). Glucocorticoid suppressible hyperaldosteronism results from hybrid genes created by unequal cross-over between *CYP11B1* and *CYP11B2*. *Proc. Natl. Acad. Sci. USA* **89**:8327-8331.

Pascoe L., Curnow, K.M., Slutsker L., Rösler, A. & White, P.C. (1992b). Mutations in the human CYP11B2 (aldosterone synthase) gene causing corticosterone methyloxidase II deficiency. *Proc. Natl. Acad. Sci. USA* **89**:4996-5000.

Pedersen, R.C. & Brownie, A.C. (1987). Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumour. *Science* **236**:188-190.

Pendurthi, U.R., Lamb, J.G., Nguyen, N., Johnson, E.F. & Tukey, R.H. (1990). Characterization of the *CYP2C5* gene in 21L III/J rabbits. *J. Biol. Chem.* **265**:14662-14668.

Peterson, J.A., White, R.E., Yasukochi, Y., Coomes, M.L., O'Keefe, D.H., Ebel, B.S.S., Masters, S.S., Ballou, D.P. & Coon, M.J. (1977). Evidence that purified liver microsomal cytochrome P-450 is a one-electron acceptor. *J. Biol. Chem.* **252**:4431-4434.

Petruska, J. & Goodman, M.F, (1985). Influence of neighbouring bases on DNA polymerase insertion and proofreading fidelity. *J. Biol. Chem.* **260**:7533-7539.

Pfanner, N., Hartl, F.-U., Guiard, B. & Neupert, W. (1987). Mitochondrial precursor proteins are imported through a hydrophilic membrane environment. *Eur. J. Biochem.* **169**:289-293.

Pheiffer, B.H. & Zimmerman, S.B. (1983). Polymer-stimulated ligation: Enhanced blunt- or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions. *Nucl. Acids Res.* **11**:7853-7871.

Picado-Leonard, J. & Miller, W.L. (1987). Cloning and sequence of the human gene for P450c17 (steroid 17α-hydroxylase/17,20 lyase): Similarity to the gene for P450c21. *DNA* **6**:439-448.

Pollack, J.D. (1979). In *The Mycoplasmas*, vol. 1 (Barlie, M.F. & Razin, S. eds.) pp187-211, Academic Press, New York, USA.

Poulos, T.L., Freer, S.T., Alden, R.A., Edwards, S.L., Skoglund, U., Takio, K., Eriksson, B., Xuong, N., Yonetani, T. & Kraut, J. (1980). The crystal structure of cytochrome c peroxidase. *J. Biol. Chem.* **255**:575-580.

Poulos, T.L., Finzel, B.C., Gunsalus, I.C., Wagner, G.C. & Kraut, J. (1985). The 2.6Å crystal structure of *Pseudomonas putida* cytochrome. *J. Biol. Chem.* **260**:16122-16130.

Poulos, T.L. & Raag, R. (1992). Cytochrome P450cam: crystallography, oxygen activation, and electron transfer. *FASEB J.* **6**:674-679.

Raag, R., Martinis, S.A., Sligar, S.G. & Poulos, T.L. (1991). Crystal structure of cytochrome P-450_{cam} active site mutant Thr252Ala. *Biochemistry* **30**:11420-11429.

Rassow, J., Guiard, B., Wienhues, U., Herzog, V., Hartl, F-U, Neupert, W. (1989). Translocation arrest by reversible folding of a precursor protein imported into mitochondria. A means to quantitate translocation contact sites. *J. Cell. Biol.* **109**:1421-1428.

Ravichandran, K.G., Boddupalli, S.S., Haseman, C.A., Peterson, J.A. & Deisenhoffer, J. (1993). Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. *Science* **261**:731-736.

Rhéaume, E., Simard, J., Morel, Y., Mebarki, F., Zachmann, M., Forest, M., New, M.I. & Labrie, F. (1992). Congenital adrenal hyperplasia due to point mutations in the type II 3ß-hydroxysteroid dehydrogenase gene. *Nature Genet.* **1**:239-245.

Riess, O., Noerremoelle, A., Collins, C., Mah, D., Weber, B. & Hayden, M.R. (1992). Exclusion of DNA changes in the ß-subunit of the c-GMP phosphodiesterase gene as the cause for Huntington's disease. *Nature Genet.* **1**:104-113.

Roise, D., Horvath, S.D., Tomich, J.M., Richards, J.H. & Schatz, G. (1986). A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artifical phospholipid bilayers. *EMBO J.* **5**:1327-1334.

Rose, M.D., Misra, L.M. & Vogel, J.P. (1989). *KAR2*, a karyogamy gene, is the yeast homolog of the mammalian BiP/ GRP78 gene. *Cell* **57**:1211-1220.

Rose, S. (1979). In *The Chemistry of Life*, 2nd. edition, Sources of energy; p.158. Penguin Books Ltd, Harmondsworth, England.

Rösler, A., Weshler,N., Lieberman, E., Hochberg, Z., Weidenfeld, J., Sack, J. and Chemke J. (1988). 11ß-hydroxylase deficiency congenital adrenal hyperplasia: update of prenatal diagnosis. *J. Clin. Endocrinol. Metab.* **66**:830-838.

Rumsby, G., Carroll, M.C., Porter, R.R., Grants, D.P. & Hjelm, M. (1986). Deletion of the steroid 21-hydroxylase and complement C4 genes in congenital adrenal hyperplasia. *J. Med. Genet.* **23**:204-209.

Rumsby, G., Skinner, C., Lee, H.A. & Honour, J.W. (1992a). Combined 17α-hydroxylase/ 17,20 lyase deficiency caused by heterozygous stop codons in the cytochrome P450 17α-hydroxylase gene. *Clin. Endocrinol.* **39**:483-485.

Rumsby, G., Skinner, C. and Honour J.W. (1992b). Genetic analysis of the steroid 21-hydroxylase gene following *in vitro* amplification of genomic DNA. *J. Steroid. Biochem.* **41**:827-829.

Ryan, K.J. & Engel, L.L. (1957). Hydroxylation of steroids at carbon 21. *J.Biol.Chem.* **225**:103-114.

Sakaguchi, M., Mihara, K. & Sato, R. (1984). Signal recognition particle is required for co-translational insertion of cytochrome P-450 into microsomal membranes. *Proc. Natl. Acad. Sci. USA* **81**:3361-3364.

Sakaki, T., Shibata, M., Yabusaki, Y., Murakami, H. & Ohkawa, H. (1989). Expression of bovine cytochrome P450c17 cDNA in *Saccharomyces cerevisiae. DNA* **8**:409-418.

Salemme, F.R. (1977). Structure and function of cytochromes C. *Annu. Rev. Biochem.* **46**:299-329.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989a). In *Molecular cloning: a laboratory manual* 2nd Edition. (1989), Volume 3 Appendix B.4. Coldspring Harbor Laboratory Press, New York, USA.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989b). In *Molecular cloning: a laboratory manual* 2nd Edition. (1989), Volume 3 Appendix C.1. Coldspring Harbor Laboratory Press, New York, USA.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989c). In *Molecular cloning: a laboratory manual* 2nd Edition. (1989), Volume 1 Page 2.100. Coldspring Harbor Laboratory Press, New York, USA.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989d). In *Molecular cloning: a laboratory manual* 2nd Edition. (1989), Volume 1 Page 2.108. Coldspring Harbor Laboratory Press, New York, USA.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**:5463-5467.

Sarkar, G., Yoon, H.S. & Sommer, S.S. (1992). Screening for mutations by RNA single-strand conformation polymorphisms (rSSCP): Comparison with DNA-SSCP. *Nucl. Acids Res.* **20**:871-878.

Sasaki, K., Yamano, Y., Bardhan, S., Iwai, N., Murray, J.J., Hasegawa, M., Matsuda, Y. & Inagami, T. (1991). Cloning and expression of a complementary DNA encoding a bovine adrenal angiotensin II type-1 receptor. *Nature* **351**:230-232.

Savitz, A.J. & Meyer, D.I. (1990). Identification of a ribosome receptor in the rough endoplasmic reticulum. *Nature* **346**:540-544.

Scalenghe, F., Turco, E., Edström, J.E., Pirotta, V. & Melli, M. (1981). Microdisection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. *Chromosoma* **82**:205-216.

Scallen, T.J., Noland, B.J., Gavey, K.L., Bass, N.M., Ockner, R.K., Chanderbhan, R. & Vahouny, G.V. (1985). Sterol carrier protein 2 and fatty acid-binding protein: separate and distinct physiological functions. *J. Biol. Chem.* **260**:4733-4739.

Schneider, M.A. (1991). Adrenal and extra-adrenal production of 11-deoxycorticosterone. PhD Thesis. pp181-197. University of London.

Serizawa, N. & Matsuoka, T. (1991). A two component-type cytochrome P-450 monooxygenase system in a prokaryote that catalyzes hydroxylation of ML-236B to pravastatin, a tissue-selective inhibitor of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Biochim. Biophys. Acta.* **1084**:35-40.

Sheffield, V.C., Cox, D.R. & Myers, R.M., (1989). Attachment of a 40-base-pair G+C rich sequence to genomic DNA fragments by the polymerase chain reaction results in improved detection of single base changes. *Proc. Natl. Acad. Sci. USA* **86**:232-236.

Sheffield, V.C., Beck, J.S., Lidral, A., Nichols, B., Cousineau, A. & Stone, E.M. (1992). Detection of polymorphisms within gene sequences by GC-clamped denaturing gradient gel electrophoresis. *Am. J. Hum. Genet.* **50**:567-575.

Sheffield, V.C., Beck, J.S., Kwitck, A.E., Sandstrom, D.W. & Stone, E.M. (1993). The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substituions. *Genomics* **16**:325-332.

Simpson, E.R. & Waterman, M.R. (1988). Regulation of the synthesis of steroidogenic enzymes in adrenal cortical cells by ACTH. *Annu. Rev. Physiol.* **50**:427-440.

Sklar, C.A., Kaplan, S.L. & Grumbach, M.M. (1980). Evidence for dissociation between adrenarche and gonadarche: Studies in patients with idiopathic precocious puberty, gonadal dysgenesis, isolated gonadotrophin deficiency, and constitutionally delayed growth and adolescence. *J. Clin. Endocrinol. Metab.* **51**:548-556.

Skinner, C.A. & Rumsby, G. (1994). Steroid 11ß-hydroxylase deficiency caused by a five base pair duplication in the *CYP11B1* gene. *Hum. Mol. Genet.* **3**:377-378.

Sparkes, R.S., Klisak, I. & Miller, W.L. (1991). Regional mapping of genes encoding steroidogenic enzymes: P450scc to 15q23-24, adrenodoxin to 11q22, adrenodoxin reductase to 17q24-q25 and P450c17 to 10q24-q25. *DNA Cell Biol.* **10**:359-365.

Speiser, P.W., Agdere, L., Ueshiba, H., White, P.C. & New, M.I. (1991). Aldosterone synthesis in patients with salt-wasting congenital adrenal hyperplasia (21-hydroxylase deficiency) and complete absence of adrenal 21-hydroxylase (P450c21). *N. Engl. J. Med.* **324**:145-149.

Spoudeas, H.A., Slater, J.D.H., Rumsby, G., Honour, J.W. & Brook, G.D. (1993). Deoxycorticosterone, 11ß-hydroxylase and the adrenal cortex. *Clin. Endocrinol.* **39**:245-251.

Suzuki, Y. Orita, M., Shiraishi, M., Hayashi, K. & Sekiya, T. (1990). Detection of *ras* gene mutations in human lung cancers by single-strand conformation polymorphism analysis of polymerase chain reaction products. *Oncogene* **5**:1037-1043.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. & Stahl, F.W. (1983). The double-strand-break repair model for recombination. *Cell* **33**:25-35.

Taubman, M.B., Berk, B.C., Izumo, S., Tsuda, T., Alexander, R.W. & Nadal-Ginard, B. (1989). Angiotensin II induces c-fos mRNA in aortic smooth muscle. Role of Ca^{2+} mobilization and protein kinase C activation. *J. Biol. Chem.* **264**:526-530.

Ten Eyck, L.F. (1979). In *The Porphyrins* (Dolphin, D. ed.) Vol. 7, Part B, pp.445-472. Academic Press, New York.

Tsubaki, M., Hiwatasi, A. & Ichikawa, Y. (1989). Conformational change of the heme moiety of the ferrous cytochrome P-450_{scc}-phenyl isocyanide complex on binding of reduced adrenodoxin. *Biochem.* **28**: 9777-9784.

Ulick, S., Wang, J.Z. & Morton, H. (1992). The biochemical phenotypes of two inborn errors in the biosynthesis of aldosterone. *J. Clin. Endocrinol. Metab.* **72**:1415-1420.

Ullmann, A., Jacob, F. & Monod, J. (1967). Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the ß-galactosidase structural gene of *Escherichia coli*. *J. Mol. Biol.* **24**:339-343.

Veldhuis, J.D., Kulin, H.E., Santen, R.J., Wilson, T.E. & Melby, J.C. (1980). Inborn error in the terminal step of aldosterone biosynthesis. *N. Engl. J. Med.* **303**:117-121.

Vermilion, J.L. & Coon, M.J. (1978). Purified liver microsomal NADPH-cytochrome P-450 reductase. *J. Biol. Chem.* **253**:2694-2704.

Verner, K. & Schatz, G. (1988). Protein translocation across membranes. *Science* **241**:1307-1313.

Vieira, J. & Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259-268.

von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J.* **5**:3021-3027.

Vos, J.-M. & Wauthier, E.L. (1991). Differential introduction of DNA damage and repair in mammalian genes transcribed by RNA polymerase I and II. *Mol. Cell. Biol.* **11**:2245-2252.

Voutilainen, R., Tapainen, J., Chung, B.-C., Matteson, K.J. & Miller, W.L. (1986). Hormonal regulation of P450scc (20,22 desmolase) and P450c17 (17α-hydroxylase/17,20lyase) in cultured human granulosa cells. *J. Clin. Endocrinol. Metab.* **63**:202-207.

Voutilainen, R. & Miller, W.L. (1987). Coordinate tropic hormone regulation of mRNA's for insulin-like growth factor II and the cholesterol side-chain cleavage enzyme, P450scc, in human steroidogenic tissues. *Proc. Natl. Acad. Sci. USA* **84**:1590-1594.

Wada, A., Okamoto, M., Nonaka, Y. & Yamano, T. (1984). Aldosterone biosynthesis by a reconstituted P45011ß system. *Biochem. Biophys. Res. Commun.* **119**:365-371.

Wagner, M.J., Ge, Y., Siciliano, M. & Wells, D.E. (1991). A hybrid cell mapping panel for regional localization of probes to human chromosome 8. *Genomics* **10**:114-125.

Wassarman, D.A. & Steitz, J.A. (1993). Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. *Science* **257**:1918-1925.

Werkmeister, J.W., New, M.I., Dupont, B. & White, P.C. (1986). Frequent deletion and duplication of the steroid 21-hydroxylase genes. *Am. J. Hum. Genet.* **39**:461-463.

White, K.A. & Marletta, M.A. (1992). Nitric oxide synthase is a cytochrome P-450 type hemoprotein. *Biochemistry* **31**:6627-6631.

White, M.B., Carvalho, M., Derse, D., O'Brien, S.J. & Dean, M. (1992). Detecting single base substitutions as heteroduplex polymorphisms. *Genomics* **12**:301-306.

White, P.C., New, M.I. & Dupont, B. (1984). Cloning and expression of cDNA encoding a bovine adrenal cytochrome P-450 specific for steroid 21-hydroxylation. *Proc. Natl. Acad. Sci. USA* **81**:1986-1990.

White, P.C., New, M.I. & Dupont, B. (1986). Structure of the human 21-hydroxylase genes. *Proc. Natl. Acad. Sci. USA* **83**:5111-5115.

White, P.C., Dupont, B., New,M.I., Leiberman, E., Hochberg, Z. & Rösler, A. (1991). A mutation in CYP11B1 (Arg-448→His) associated with steroid 11ß-hydroxylase deficiency in Jews of Moroccan origin. *J. Clin. Invest.* **87**:1664-1667.

White, P.C., & Pascoe, L. (1992). Disorders of steroid 11ß-hydroxylase isozymes. *Trends in Endocrinology and Metabolism* **3**:229-234.

Wickner, W.T. & Lodish, H.A. (1985). Multiple mechanisms of protein insertion into and across membranes. *Science* **230**:400-407.

Winship, P.R. (1989). An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucl. Acids Res.* **17**:1266.

Wong, M., Rice, D.A., Parker, K.L. & Schimmer, B.P. (1989). The roles of cAMP and cAMP-dependent protein kinase in the expression of cholesterol side chain cleavage and steroid 11ß-hydroxylase genes in mouse adrenocortical tumour cells. *J. Biol. Chem.* **264**:12867-12871.

Yamano, S., Aoyama, T., McBride, O.W., Hardwick, J.P., Gelboin, H.V. & Gonzalez, F.J. (1989). Human NADPH-P450 oxidoreductase: Complementary DNA cloning, sequence, vaccinia virus-mediated expression, and localization of the CYPOR gene to chromosome 7. *Mol. Pharmacol.* **35**:83-88.

Yanagibashi, K. & Hall, P.F. (1986). Role of electron transport in the regulation of the lyase activity of C21 side-chain cleavage P-450 from porcine adrenal and testicular microsomes. *J. Biol. Chem.* **261**:8429-8433.

Yanagibashi, K., Haniu, M., Shively, J.E., Shen, W.H. & Hall, P. (1986). The synthesis of aldosterone by the adrenal cortex: two zones (fasciculata and glomerulosa) possess one enzyme for 11ß-,18-hydroxylation and aldehyde synthesis. *J. Biol.Chem.* **261**:3556-3562.

Yanase, T., Simpson, E.R. & Waterman, M.R. (1991). 17α-hydroxylase/17,20-lyase deficiency: from clinical investigation to molecular definition. *Endocr. Rev.* **12**:91-96.

Yanisch-Perron, C., Vieira, J. & Messing J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103-119.

Yasukochi, Y. & Masters, B.S.S. (1976). Some properties of a detergent-solubilized NADPH-cytochrome c (Cytochrome P-450) reductase purified by biospecific chromatography. *J. Biol. Chem.* **251**:5337-5344.

Zachmann, M., Vollmin, J.A., Hamilton, W. & Pader, A. (1972). Steroid 17,20 desmolase deficiency: A new cause of male pseudohermaphroditism. *Clin. Endocrinol.* **1**:369-385.

Zachmann, M., Tassinari, D. & Prader, A. (1983). Clinical and biochemical variability in congenital adrenal hyperplasia due to 11ß-hydroxylase deficiency. *J. Clin. Endocrinol. Metab.* **56**:222-229.

Zanger, U.M., Kagawa, N., Lund, J. & Waterman, M.R. (1992). Distinct biochemical mechanisms for cAMP-dependent transcription of *CYP17* and *CYP21*. *FASEB J.* **6**:719-723.

Zuber, M.X., Simpson, E.R. & Waterman, M.R. (1986). Expression of bovine 17α-hydroxylase cytochrome P450 cDNA in non-steroidogenic (COS 1) cells. *Science* **234**:1258-1261.

Addenda:

Abramovitz, M., Laplante Branchaud, C. & Pearson-Murphy, B.E. (1982). Cortisol-cortisone interconversion in human fetal lung: contrasting results using explant and monolayer cultures suggests that 11ß-hydroxysteroid dehydrogenase (EC 1.1.1.146) comprises two enzymes. *J Clin Endocrinol Metab.* **54**:563-568.

Edwards, C.R.W., Stewart, P.M., Burt, D., Brett, L., McIntyre, M.A., Sutanto, W.S., de Kloet, E.R. & Monder, C. (1988). Localization of 11ß-hydroxysteroid dehydrogenase- tissue specific protector of the mineralocorticoid receptor. *Lancet* **ii**, 986-989.

Griffing, G.T., Dale, S.L., Holbrook, M.M. & Melby, J.C. (1983). 19-nor-deoxycorticosterone excretion in primaryaldosteronism and low renin hypertension. *J Clin Endocrinol Metab.* **56**:218-221.

Lakshmi, V. & Monder, C. (1985). Evidence for independent 11-oxidase and 11-reductase activities of 11ß-hydroxysteroid dehydrogenase: enzyme latency, phase transitions and lipid requirements. *Endocrinology* **116**:552-560.

Lakshmi, V. & Monder, C. (1988). Purification and characterization of the corticosteroid 11ß-hydroxysteroid dehydrogenase component of the rat liver 11ß-hydroxysteroid dehydrogenase complex. *Endocrinology* **123**:2390-2398.

Phillips, D.M., Lakshmi, V. & Monder, C. (1989). Corticosteroid 11ß-hydrogenase in rat testis. *Endocrinology* **125**:205-216.