Molecular analysis of the human 3β-hydroxysteroid dehydrogenase Δ5/Δ4 isomerase gene family

A thesis submitted for the degree of
Doctor of Philosophy
at the
University of Glasgow

by

Martin William McBride

Division of Molecular Genetics
Institute of Biomedical and Life Sciences
Anderson College
Glasgow

November 1996
This PhD thesis is dedicated to my mum and dad, 
my brother and sisters Robert, Clare and Alison and finally to 
Tracey, for all their love, support and encouragement
The research reported in this thesis is my own original work, except where otherwise stated and has not been submitted for any other degree.
Table of contents

Table of Contents iv
Acknowledgements viii
Summary ix

Chapter 1
General introduction

1.1 Principles of hormone action. 1
1.2 Steroid hormones. 2
1.3 Steroid hormone biosynthetic pathways. 4
1.4 3β-hydroxysteroid dehydrogenase Δ5/Δ4 isomerase (3β-HSD). 6
1.5 Steroidogenesis in fetal tissues. 6
1.6 Adult adrenal steroidogenesis. 8
1.7 Testicular and Ovarian steroidogenesis. 11
1.8 Congenital adrenal hyperplasia (CAH). 13
1.9 Molecular biology of 3β-hydroxysteroid dehydrogenase Δ5/Δ4 isomerase (3β-HSD). 19
1.10 Structure-function predictions of 3β-HSD isozymes. 20
1.11 Tissue distribution of 3β-HSD activity in humans. 22
1.12 Subcellular localisation of 3β-HSD. 23
1.13 Background and objectives of project. 24

Chapter 2
Materials and Methods

2.1 Random priming of DNA. 27
2.2 PCR primer sequences. 27
2.3 Screening of human placental expression library λ gt 11 and PCR reaction conditions. 28
2.4 Agarose gel electrophoresis. 29
2.5 Mammalian cell culture. 29
2.6 Transfection of mammalian cells using CaPO4. 30
2.7 Selection of Stable CHO cell lines. 30
2.8 Immunostaining of transfected cells. 30
2.9 Fusion proteins. 31
2.10 SDS-polyacrylamide gel electrophoresis and western immunoblot analysis. 31
2.11 Peptides. 32
2.12 Peptide Competition experiments. 32
2.13 Screening of human genomic libraries in λ gem11 and PCR reaction 32

iv
conditions.

2.14 Polymerase chain reaction (PCR).
2.15 Denaturing gradient gel electrophoresis.
2.16 Confirmation of putative clones by PCR.
2.17 pT7 Blue T-vector cloning of PCR products.
2.18 Preparation of plasmid DNA.
2.19 Dideoxy DNA sequencing.
2.20 Direct sequencing of PCR products.
2.21 Bacteriophage lambda DNA.
2.22 Bidirectional Southern blot analysis of lambda DNA.
2.23 Phylogenetic analysis.

Chapter 3
Cloning and expression of the cDNA clone B3 (P10), and it's identification as 3β-hydroxysteroid $\Delta^5 / \Delta^4$ isomerase type I (3β-HSD type I).

3.1 Introduction 37

Results
3.2 Isolation of B3, a 1700bp placental cDNA clone. 38
3.3 B3 (3β-HSD type I) PCR cloning into the eukaryotic expression vector pcDNAI/Neo. 40
3.4 Transient expression of pCDNeo/B3 in COS-7 cells and immunostaining with FDO 161G monoclonal antibody. 43
3.5 FDO 26G staining of CHO cells stably transfected with pcDNeo/B3. 45
3.6 Evidence for Bgl II polymorphism from clones isolated in the placental library screen. 46
3.7 DGGE analysis of positive clones isolated from the placental library. 47

Chapter 4
Mapping the monoclonal antibody FDO 26G epitope to the carboxy-terminus of B3 (3β-HSD type I).

4.1 Introduction. 50

Results
4.2 Mapping the FDO 26G epitope using bacterial lacZ fusion genes. 50
4.2.1 Constructs pURMM1 and pURMM2 (Residues 1-241 and 210-373). 52
4.2.2 Construct pURMM3 (Residues 263-373). 55
4.2.3 pURMM4 and 5 constructs. (3β-HSD type I residues 241-285, and 289-373). 58
4.3 Western immunoblot analysis of inclusion body preparations from constructs pURMM1-5. 62
4.4 Peptide mapping experiments. 62
4.5 Summary of lacZ gene fusion and oligopeptide mapping data, localising the FDO 26G epitope to residues 354-366 of the B3 (3β-HSD type I) protein. 64
4.6 Effect of phosphorylation of S-359 on the binding of FDO 26G to peptide 354-366. 65

Chapter 5
Characterisation of clone 24-4, a new member of the human 3β-HSD gene family and its identification as a presumptive unprocessed pseudogene (3β-HSDψ1).

5.1 Introduction 66

Results

5.2 Isolation of the clone 24-4 from the Phe family genomic library and its identification as the ψ gene. 69
5.3 Isolation of clone 1-1 from a control genomic library. 71
5.4 Southern analysis of of bacteriophage clones 24-4 and 1-1. 73
5.5 Subcloning clone 24-4 putative exons into pUC18. 76
5.6 Subcloning clone 1-1 putative exons into pUC18. 76
5.7 Sequencing 24-4 and 1-1 subclones. 77
5.8 Clone 24-4 and clone 1-1 identification as 3β-HSDψ1. 83

Chapter 6
Characterisation of further novel members of the human 3β-HSD gene family.

6.1 Introduction. 84

Results

6.2 Analysis of genomic clones by exon 4B (G,H) PCR amplification and DGGE analysis. 86
6.3 Summary DGGE gel of exon 4B analysis. 89
6.4 Characterisation of novel clones identified from two human genomic libraries by exon 4B DGGE analysis. 90
6.4.1 Clone 19-4. DGGE classification α. 90
6.4.2 Clone 6-5. DGGE classification γ. 95
6.4.3 Clone 1-2. DGGE exon 4B classification ρ fragment. 96
6.4.4 Clone 2-1. DGGE exon 4B classification T fragment. 96
6.5 Analysis of genomic clones by exon 3 (A, C) PCR amplification and DGGE analysis. 97
6.6 Identification of further members of the human 3β-HSD gene family by PCR amplification followed by direct sequencing of PCR products. 98
   6.6.1 Clone 4-3 98
   6.6.2 Clone 8-3 98
6.7 Summary table of clones and classifications. 101
6.8 Sequence alignment (Figure 6.13). 103
6.9 Sequence divergence and phylogenetic analysis of the human 3β-HSD gene family. 110

Chapter 7
General discussion

7.1 Mapping the monoclonal antibody FDO 26G epitope. 113
7.2 Investigation into the extent of the human 3β-HSD gene family. 116

Chapter 8
Bibliography 123
Acknowledgements

Firstly I would like to thank Roger Sutcliffe for his support and advice during the completion of this thesis. Thanks also to Elizabeth Boyd and Norma Morrison, who were responsible for mapping the genomic clones to human chromosomes. In addition I am very grateful to 'Auntie' Margaret, Mary Burke and Mary Gardner whose help and friendship have meant a great deal to me. Thanks also to the ladies of the prep room who have provided an endless supply of entertainment.

I have appreciated the assistance provided by members of my group both past and present; Arvind, Sirous, Alan, Gertrudes, Simon, Sandra and more recently Alison and Nicola. Very special thanks to Richard Wilson, Marshall, Kevin and Jonathan for many helpful and morale boosting comments. Thanks especially to many members of the 'sixth floor', including Dave, Jen, and Richard.

Thanks to the other members of the Genetics department that have made mass gatherings worthwhile. I would especially like to thank Colin, Gareth and Leslie for their support, love and friendship.
Summary

3-beta hydroxysteroid dehydrogenase $\Delta^5/\Delta^4$ isomerase (3B-HSD) catalyses the oxidative conversion of $\Delta^5$-3beta-hydroxysteroids to their corresponding $\Delta^4$-3-ketosteroid configuration, and is essential for the biosynthesis of all classes of steroid hormone; mineralocorticoids, glucocorticoids, progestins, androgens and oestrogens. In humans two homologous linked, genes, HSD3B1 an HSD3B2, encoding 3B-HSD type I and type II enzymes, have been isolated and mapped to the human chromosomal region 1p13.1. RNase protection assays have localised 3B-HSD type I transcripts principally in placenta and skin, while 3B-HSD type II transcripts are almost exclusively found in the adrenals and gonads. The two forms are 93.5% identical in primary structure and antibodies raised against purified human placental protein react immunocytochemically with 3B-HSD in all four tissues. Such tissue specificity is shown by monoclonal antibody FDO 161G which reacts with 3B-HSD, and was originally raised against human syncytiotrophoblast cells from the placenta. However, a further mouse monoclonal antibody, FDO 26G, raised against purified 3B-HSD from placental villous tissue showed more restrictive tissue reactivity. Immunocytochemically, FDO 26G reacted with 3B-HSD in villous syncytiotrophoblast and adrenals, but reacted weakly with extravillous trophoblast and Leydig cells of the testis. These experimental observations provide evidence for the expression of a further isoform of 3B-HSD exists in extravillous trophoblast and adult Leydig cells; either a post-translational modification within the region of monoclonal antibody binding (the epitope) or the expression of as yet unidentified 3B-HSD gene that encodes a different amino acid sequence over the FDO 26G epitope.

Using a combination of lac Z fusion polypeptides and synthetic peptides, the FDO 26G epitope was located to residues 354-366 at the carboxy terminal end of 3B-HSD type I, an amino sequence that is identical in the type I and type II forms of the enzyme. This epitope contains a consensus for caesin kinase II phosphorylation, with serine 359 as the candidate target of phosphorylation. This suggested that the lack of reactivity of FDO 26G in certain trophoblast cells of the placenta might be due to phosphorylation of serine 359. Peptide 354-366 was synthesised with phosphoserine at residue 359. FDO 26G reactivity to the phosphopeptide and unphosphorylated peptide was compared. FDO 26G bound the phospho-peptide at least as strongly as the unphosphorylated peptide. It was concluded that phosphorylation of serine 359 was not responsible for lack of FDO 26G reactivity in placental trophoblast populations.

A number of 3B-HSD cDNAs and genes have been cloned from a variety of vertebrate species. In mouse six homologous 3B-HSD cDNAs have been isolated and shown to be
expressed tissue- and sex-specifically. This gene family clustered on mouse chromosome 3 includes two functionally distinct groups of proteins; enzymes that catalyse the oxidative conversion of Δ5-3β-hydroxysteroids (characterised by an aspartic acid at residue 36) and enzymes that exclusively act as 3-ketosteroid reductases (characterised by a tyrosine (rat) or phenylalanine (mouse) at residue 36).

Several bands of hybridisation are detected when Southern blots of human genomic DNA were probed with HSD3B1 exon specific probes suggesting that 3β-HSD-like sequences other than HSD3B1 and HSD3B2 exist in the human genome. In addition when screening amplified segments of patient DNA for mutations in 3β-HSD types I and II, it is relatively common to amplify novel but closely related 3β-HSD sequences.

To estimate the size of the 3β-HSD gene family two human genomic λgem 11 libraries were screened with 3β-HSD type I cDNA under non-stringent conditions. From a total of $1.4 \times 10^6$ clones, fifty-seven positive clones hybridised reproducibly through to the second stage of screening. These clones were tested as templates for PCR using primers that were known to amplify homologous members of the gene family (from mutation screening). PCR products were classified into groups depending on their mobility on denaturing gradient electrophoresis gels. Representatives of each group were sequenced. The remaining PCR-amplified products were directly sequenced. A total of seven distinct 3β-HSD gene family members, including HSD3B1 and HSD3B2, were identified and mapped to the human chromosomal region 1p13. All of the newly identified genes contained 3β-HSD-like exon and intron sequences indicating that these clones were not processed pseudogenes. Novel 3β-HSD sequences were subcloned and sequenced from each of the new members of the family. The presence of frameshift and in some cases, missense mutations associated with 3β-HSD type II deficiency suggests that these genes are not expressed as functional 3β-HSD enzymes. There was no evidence of a human 3β-HSD ketosteroid reductase enzyme.

Based on alignments of the predicted cDNAs, from each of the novel genes, human 3β-HSD type I and II, macaque, bovine and the rodent 3β-HSD gene families, a phylogenetic relationship was generated using the maximum parsimony method. The analysis suggests that several gene duplication and events have occurred independently, in the evolution of the 3β-HSD gene family in different species. Since primates and rodents diverged, an ancestral 3β-HSD gene(s) has duplicated a number of times in the primate lineage and several gene duplications events have taken place in the rodent lineage. Interestingly, the amplification of the 3β-HSD gene family in rodents leads to many 3β-HSD genes expressed tissue- and sex-specifically, whereas in humans, there appear to be only two genes expressed in a tissue specific manner.
Chapter 1

General Introduction

1.1 Principles of hormone action.

Regulation of metabolic processes and many other aspects of multicellular organism function are controlled by hormones (chemical mediators) of the endocrine, paracrine, autocrine, intracrine and neurotransmitter systems. Hormones maintain the internal environment in a state of homeostasis, and enable it to react to changes in the external environment. Even unicellular organisms produce molecules that are very similar to the hormonal peptides of vertebrates, suggesting that the biochemistry of co-ordinate cell to cell communication may have its origins in simple unicellular organisms (LeRoith et al 1980; Roth et al 1982).

Endocrinology is the broad discipline of the study of the action of hormones and the organs in which they are synthesised and secreted. Hormones can be classified in many ways, including by the cells they target (Figure 1.1).

![Diagram of hormone types](Image)

Figure 1.1. Schematic diagram of endocrine, paracrine, autocrine, and intracrine action. Classically, endocrine hormones are synthesised and secreted in specialised glands for release into the general circulation and transport to distant target cells. In addition, hormones released from one cell can influence neighbouring cells (paracrine activity) or can exert a positive or negative effect on the secretory cell itself (autocrine activity). Intracrine activity describes the local formation of active hormones which influence the cells in which they are synthesised without release into the extracellular compartment.

Classically hormones were divided into four groups; peptides, steroids, idiothyronines and catecholamines. However, new classes of ligands have been added to the list including prostaglandins, amino acids and lipids. In each case, these molecules interact with target cells by binding with receptor proteins which recognise the hormone selectively.
conferring specificity of response. Despite the multiple chemical classes of hormones, hormone receptors fall into two general classes. Charged molecules such as peptide hormones and neurotransmitters interact with cell surface plasma membrane receptors and act through stimulation of second messenger pathways (Dohlman et al 1987). Uncharged (lipophillic) molecules such as steroid and thyroid hormones diffuse into the cell and bind intracellular receptor proteins with high affinity (Evans et al 1988), which may be found within the cytoplasmic or nuclear compartments. The hormone-receptor complex binds to specific DNA sequences, the hormone response element that is positioned upstream of the transcription initiation site. As a result, mRNA and subsequent protein synthesis are affected.

1.2 Steroid hormones

Steroid hormones control many metabolic and biosynthetic pathways, in almost all organs and tissues. All steroid hormones are derivatives of cholesterol and are structurally similar, containing the cyclopentanoperhydrophenanthine nucleus (Figure 1.2).

Various side chains and degrees of saturation confer specificity to the different hormone molecules by their action on one or more specific hormone receptors. As regulators of physiological processes, they are categorised in five main groups; mineralocorticoids, glucocorticoids, progestins, androgens and oestrogens.

Mineralocorticoids

Aldosterone is the principal mineralocorticoid and is normally produced exclusively in the zona glomerulosa of the adrenal cortex (Figure 1.3). Other steroids in the aldosterone biosynthetic pathway also have mineralocorticoid activity, and act by binding the Type I (mineralocorticoid) receptor (Arriza et al 1987). Mineralocorticoids regulate Na\(^+\) and K\(^+\)
transport in epithelial cells such as the proximal and distal tubules of the kidney, salivary and sweat glands and the gastrointestinal tract. These cells respond in a similar manner to mineralocorticoids, by increasing salt (Na\(^+\)) absorption from various tubules or ducts back into the body while excreting K\(^+\) ions (Morris et al 1981). Patients with mineralocorticoid deficiency develop weight loss, hyperkalemia and are hypotensive, whereas patients with mineralocorticoid excess develop hypernatremia and are hypertensive.

*Glucocorticoids*

Glucocorticoids, mainly cortisol in humans, influence a wide variety of metabolic processes in many organs and tissues of the body (Schulster, D. 1976). They have important effects on carbohydrate metabolism, raising the concentration of glucose in the blood by increasing the synthesis of a number of enzymes in gluconeogenesis (Cidlowski et al 1982). The glucocorticoids also alter protein metabolism causing breakdown of muscle tissue, mobilising an alternative energy source to glucose. Other effects of glucocorticoids include suppression of DNA synthesis and inhibition of the inflammatory response to infection or injury. Increased secretion of cortisol also occurs as a response to stress.

*Androgens*

The tissues and organs that are regulated by androgens are diverse. Androgens stimulate male sexual differentiation, and maturation of the reproductive organs *in utero*. They also promote development of male secondary sexual characteristics and behaviour. The two most important androgens are testosterone and 5α-dihydrotestosterone (Figure 1.3 and sections 1.5-1.7).

*Oestrogens and progestins*

In both male and female, oestrogens are derived from either the adrenal gland or the gonad. The enzyme P450arom (aromatase) converts androgens to oestrogens and is present in many tissues of the body (Figure 1.3). Progesterone is derived from the conversion of pregnenolone by 3β-hydroxysteroid dehydrogenase (Figure 1.3 and sections 1.5-1.7). Oestrogens and progesterone interact to control the growth, development and physiology of the female reproductive tract, to regulate the menstrual cycle and induce
development of female secondary sexual characteristics such as mammary glands. They are also key hormones for fertility and maintenance of pregnancy (Siiteri et al 1967).

1.3 Steroid hormone biosynthetic pathways

Steroid biosynthesis occurs in a variety of organs, including adrenal cortex, testes and ovaries, and during pregnancy the placenta develops as an additional source of steroid hormones. The principal pathways of steroid hormone biosynthesis were established over twenty years ago through advances in steroid radioisotopic and chromatographic techniques (Schulster D 1976). Cholesterol is essential for steroid hormone and bile acid biosynthesis and is available from three sources, de novo synthesis from acetate, from plasma low density plasma lipoprotein or through storage as lipid inclusions (Gwynne & Strauss 1982). Steroid hormone biosynthesis in all endocrine organs is similar. It is catalysed by six forms of cytochrome P450 (Nebert & Gonzalez 1987), and two forms of hydroxysteroid dehydrogenase, many of which utilise multiple steroid substrates and have more than one enzyme activity (Figure 1.3). Tissue-specific expression of these distinct enzymes determines the steroid hormones that are synthesised. These enzymes are present within the mitochondria and the endoplasmic reticulum; therefore intermediates in steroidogenesis must 'shuffle' between distinct organelles, a process that is not well understood. Different electron transport chains are utilised depending whether enzymatic reactions takes place in the endoplasmic reticulum or mitochondria (Miller 1988).

Steroid hormones are not stored to any great extent within endocrine cells. There are large stores of cholesterol and pituitary trophic hormones (e.g ACTH and gonadotropins) stimulate cholesterol transfer from the outer to inner mitochondrial membrane by activating the steroid acute response protein (StAR) increasing the availability of substrate cholesterol for steroidogenesis (Sugawara et al 1995). This function is in addition to and reinforces the induction effects these hormones have on specific pathways and hence contribute to feedback loops between the anterior pituitary and the end organ.
Figure 1.3. Principal pathways of steroidogenesis in the adrenal cortex and gonadal tissues. 3β-HSD utilises multiple substrates and is required for the synthesis of all classes of steroid hormone. The major steroid products are highlighted in bold as is the category of steroid hormone to which they belong, except for progesterone which is a progestin. The enzymes P450scc, P450c11B1 and P450c11B2 are located in the inner mitochondrial membrane. P450c17, P450c21, P450arom are localised to the endoplasmic reticulum. 3β-HSD activity localises to both compartments. The 17β-hydroxysteroid dehydrogenase (17β-HSD) reaction is catalysed by multiple isozymes that favour the reduction reaction in testis and ovary and the oxidation reaction in extragonadal tissues.
1.4 3β-hydroxysteroid dehydrogenase $\Delta^5/\Delta^4$ isomerase (3β-HSD)

The enzyme 3β-hydroxysteroid dehydrogenase $\Delta^5/\Delta^4$ isomerase (3β-HSD) catalyses essential early steps in the biosynthesis of all classes of biologically active steroid hormone; mineralocorticoids, glucocorticoids, progestins, androgens and oestrogens (Figure 1.3). It catalyses the oxidation of multiple 3β-hydroxy-$\Delta^5$-steroid substrates to their corresponding $\Delta^4$-3-oxo configuration utilising NAD$^+$ as cofactor (Figures 1.3 and 1.4).

In the following sections steroidogenesis in various human tissues will be described, emphasising the central role of 3β-HSD activity in the biosynthesis of steroid hormones, followed by a brief description of genetic disease caused by deficiencies in certain steroidogenic enzymes. As a large proportion of my work is concerned with the human 3β-HSD gene family, emphasis will be placed on genes encoding steroidogenic enzymes present in multiple copies in the human genome.

1.5 Steroidogenesis in fetal tissues

The biosynthesis of steroids in the placenta differs in several important aspects from the biosynthetic route in other endocrine tissues such as the adrenal cortex. The placenta is deficient in several important enzymes, such as P450c17 and enzymes necessary for de novo synthesis of cholesterol, although it has considerable capacity to aromatise androgens to oestrogens. This led to the concept of the fetal placental unit (Diczfalusy...
in which the fetus and placenta complement each other in their ability to synthesise the many different steroids produced during human pregnancy (Figure 1.5).

During human pregnancy the placenta is the major site of 3\(\beta\)-HSD expression in fetal compartments. Large amounts of progesterone are produced by the placenta and released into the maternal circulation (Milewich et al 1975). Placental 3\(\beta\)-HSD activity is also essential for the conversion of DHEA to oestrogens (Figure 1.5).

The human fetal adrenal (HFA) has two morphologically and steroidogenically distinct zones, the neocortex (20\%) and fetal (80\%) zones. The neocortex zone is the primary site for cortisol synthesis and secretion, initially utilising placental progesterone as substrate; but eventually expression of 3\(\beta\)-HSD in the neocortex zone allows synthesis of cortisol from \(\Delta^5\) steroid intermediates of fetal adrenal origin. The fetal zone is more steroidogenically active than the neocortical zone, using placental pregnenolone to synthesise dehydroepiandrosterone-sulphate (DHEA-S) and dehydroepiandrosterone (DHEA) which serve as precursors for placental oestrogen formation (Sitteri et al 1973).

Fetal testicular 3\(\beta\)-HSD enzyme activity is critical for the biosynthesis of testosterone, the major androgenic hormone that controls primary male sexual characteristics (Jost et al 1973). By the seventh week of uterine life, the fetus contains the primordia of both male (Wolffian) and female (Mullerian) genital ducts. Testosterone induces the differentiation of the Wolffian ducts into the epididymis, vas deferens, seminal vesicles and ejaculatory ducts of the male. Functional fetal testes also secrete anti-Mullerian hormone causing involution of the Mullerian ducts. Although testosterone is the principal androgen secreted, in many peripheral target tissues testosterone is converted to the more potent androgen 5\(\alpha\)-dihydrotestosterone (5\(\alpha\)-DHT), which is critical for the masculinisation of external genitalia. There is a critical period for the action of androgens, after about the twelfth week of gestation fusion of the labioscotal folds will not occur, leading to incomplete masculinisation of the male fetus and hypospadias (Hiort et al 1994). Exposure of the female fetus to androgens during this period can result in virilisation of female external genitalia (Kelnar et al 1993).
Figure 1.5. Biosynthesis of progesterone and oestrogens (oestrone, oestradiol and oestriol) from the fetal-placental unit. Placental 3β-HSD enzyme activity is required for the maintenance of pregnancy and normal development of the fetus.

1.6 Adult adrenal steroidogenesis

The adult adrenal glands are endocrinologically complex organs, situated near the anterior poles of each kidney, embedded in perirenal adipose tissue. Each adrenal gland consists of two distinct organs, one surrounding the other, with separate function and different embryological origins. The inner medulla arises from the neural crests, and secretes the catecholamines (e.g. adrenaline and noradrenaline). Stress can cause an immediate release of adrenaline which prepares the body for extraordinary physical and mental exertion.
The adrenal cortex emanates from the mesoderm at an early stage of development of the embryo. The adrenal cortex is divided morphologically and functionally into three distinct zones, zona glomerulosa (ZG), zona fasciculata (ZF) and zona reticularis (ZR). Tissue specific expression of steroid hormone biosynthetic enzymes determine the steroid hormones synthesised in different zones (Figure 1.6). The ultrastructure of adrenocortical cells is similar to that of other steroid secreting cells, including the corpus luteum of the ovary and Leydig interstitial cells of the testis. They have large amounts of smooth endoplasmic reticulum, numerous lipid inclusions (cholesterol stored as esters with fatty acids), abundant mitochondria, and a prominent Golgi complex (James 1979; The Adrenal Gland, Raven press).

The outer zona glomerulosa (ZG) constitutes about 15% of the cortex, synthesising mineralocorticoids, principally aldosterone (Figure 1.6), mainly under the control of the renin-angiotensin system (Gibbons et al 1984). The zona fasciculata (ZF) is deep, approximately 75% of the cortex, and synthesises glucocorticoids, mainly cortisol. Cortisol secretion in the ZF is primarily controlled by periodic releases of a single chain polypeptide, adrenocorticotropic hormone (ACTH). ACTH is synthesised in the anterior pituitary corticotroph cells, and is released on stimulation of the cell by the hypothalamic peptide corticotrophin-releasing factor (CRF). ACTH binds to high affinity plasma membrane receptors of adrenocortical cells, activating the adenylate cyclase system (cAMP). The principal physiological action of ACTH is to cause a rapid increase in adrenocortical steroid hormone synthesis and secretion, but it also increases growth, size and blood flow through the gland. The negative feedback loop is completed when glucocorticoids 'feedback' to the hypothalamus and the anterior pituitary to limit CRF and ACTH synthesis and secretion. Precursor intermediates in the androgen biosynthetic pathway, dehydroepiandrosterone (DHEA) and dehydroepiandrosterone-sulphate (DHEA-sulphate) are the main secretion products of the zona reticularis (ZR) which are released into the circulation.

40% of androgens in man (Moghissi et al 1984) and the majority of oestrogens in women are synthesised in extra adrenal and extragonadal target tissues from precursor steroids of adrenal origin, specifically DHEA-S and DHEA (Labrie 1988). It is therefore not surprising that the enzyme activities required for the conversion of DHEA and DHEA-S, 3β-HSD, 17β-HSD, P450 aromatase, 5α-reductase and steroid sulphatase are expressed in a large number of peripheral tissues (Milewhich et al 1991).
Figure 1.6. Steroid biosynthetic pathways in the three functionally distinct zones of the adrenal cortex (not to scale). Aldosterone and cortisol are the principal mineralocorticoid and glucocorticoid secreted by the zona glomerulosa and fasciculata. Precursor sex steroids, DHEA, DHEA-S and androstenedione are secreted mainly from the zona reticularis.
1.7 Testicular and Ovarian steroidogenesis

The testis is the male gonad. Its primary functions are the production of spermatozoa and testosterone under complex feedback controls by hypothalamic-pituitary peptides including LH and FSH. Testes consist mainly of loops of convoluted seminiferous tubules, along which spermatozoa form from primitive germ cells. Between the tubules are nests of Leydig interstitial cells which contain the necessary enzyme complement for the synthesis of testosterone. An outline of testicular steroidogenesis is given in figure 1.7.

Figure 1.7. Leydig cell steroidogenesis. Testosterone is the major androgenic steroid hormone secreted by Leydig cells. Leydig cells do not metabolise 17-OH progesterone efficiently. The potent androgenic hormone 5α-DHT is only secreted in small amounts; the major sites of formation and action of 5α-DHT are in peripheral tissues (extraadrenal and extragonadal).
The main female reproductive organs are the ovaries, fallopian tubes, uterus and the vagina. Their function is to produce the ova, ensure fertilisation, allow it to grow and develop in a suitable environment and delivered at term. The ovaries are situated on either side of the uterus. The centre is known as the medulla or stroma which consists mainly of connective tissues containing collagen fibres. Surrounding the medulla is the cortex, which consists of germinal epithelium, and follicles containing the immature ovum. The major steroid producing cells of the ovary, thecal, granulosa and corpus luteum cells contain the complete enzyme complement required for the biosynthesis of androgens, oestrogens and progestins (Figure 1.8).

Figure 1.8. Principal pathways of steroid hormone biosynthesis in the human ovary. Each cell type of the ovary contains the complete enzyme complement for the formation of oestradiol from cholesterol; however the concentrations of various enzymes differ among the cell types, and consequently hormones are synthesised in thecal and granulosa compartments as indicated. As in Leydig cells ovarian cell types do not metabolise 17-OH progesterone efficiently.
Oestrogens are produced by the ovary during follicular maturation, and stimulate glandular proliferation of the endometrium. Oestrogens also stimulate the synthesis of progesterone receptors and promote the development of female secondary characteristics later in development. Progesterone is a key hormone in fertilisation and maintenance of pregnancy. It induces deciduisation of the endometrium, inhibits uterine contractions and with oestrogens promotes the glandular development of the breast. In contrast to other steroidogenic tissues, ovarian steroidogenesis is episodic, cyclical and is precisely coordinated by a number of local factors, some of which are produced within the ovary itself. Oestrogen biosynthesis is complicated by the constantly changing population of cells as the process of follicular development, ovulation and corpus luteum formation occurs with each menstrual cycle. At maturation the Graafian follicle produces less oestrogen and more progesterone. These hormones together with gonadotrophin-releasing hormone produce a massive release of LH into the bloodstream, causing the follicle to rupture and release the ovum. The follicle now becomes the progesterone- secreting corpus luteum (yellow body) with maximum secretion around day eight post ovulation. If fertilisation does not occur, the corpus luteum secretes less progesterone and becomes the corpus albicans (white body). The endometrium collapses due to the lack of blood and the lining is lost with the menstrual flow.

The regulation of steroid hormone synthesis by enzyme localisation is complemented by substrate availability. Granulosa cells of the follicle are avascular and do not have access to low density lipoprotein (LDL) plasma substrate (Carr et al 1981). After ovulation, extensive neovasculinization of the follicle takes place, providing increased amounts of cholesterol to the luteinized granulosa cells, allowing increased production of progesterone during the luteal phase of the menstrual cycle.

1.8 Congenital adrenal hyperplasia (CAH)

A central question I investigate in results chapters 5 and 6 is the number of genes within the human 3β-HSD gene family and where I attempt to address the effect of multiple 3β-HSD genes in the generation of 3β-HSD deficiencies. To further explore this argument I would like to draw parallels between other steroidogenic deficiencies, specifically when the steroid genes include multiple members.

Congenital adrenal hyperplasia (CAH) is the most frequent cause of ambiguous genitalia and adrenal insufficiency in neonates. It results from an inherited defect in any of the enzymatic steps required to synthesise cortisol from cholesterol in the adrenal cortex, leading to a compensatory hyper secretion of ACTH by the anterior pituitary with
consequent hyperplasia of the adrenal gland (New et al 1989). Each enzyme deficiency produces abnormal levels of adrenal hormones and precursors. A defect may have clinical significance not only because cortisol is not synthesised efficiently, but because precursor steroids proximal to the block may accumulate and be shunted into other metabolic pathways, particularly that of androgen biosynthesis. Excessive androgen production results in virilization of both males and females causing hirsutism in females, penile or clitoral enlargement, precocious adrenarche and early epiphyseal closure ultimately resulting in short stature. Although any of the five different enzymes involved in cortisol biosynthesis in the adrenal cortex could be impaired (Figure 1.6), P450c21 deficiency, due to a mutation within the CYP21 gene, accounts for more than 90% of CAH patients (White P.C. 1987a and 1987b). The remainder mainly consist of P450c11 and 3ß-HSD enzyme deficiencies (Saenger et al 1995; Hermans et al 1996).

Cytochrome P450c21 deficiency

Cytochrome P450c21 deficiency is one of the most common inborn errors of metabolism in humans, impairing mineralocorticoid and glucocorticoid steroid formation exclusively in the adrenal cortex. Microsomal P450c21 converts progesterone to deoxcorticosterone and 17-hydroxyprogesterone to 11-deoxycortisol, deficiency causes accumulation of cortisol precursors which may be shunted into the adrenal androgen pathway (Figure 1.6). There are a number of recognised clinical phenotypes of P450c21 deficiency; salt wasting severe classic CAH, simple virilising, non-classic and cryptic. The severe classic P450c21 deficiency occurs 1 in 10,000-15,000 live births (Cacciari et al 1983) and is associated with genital ambiguity and varying degrees of low aldosterone levels leading to electrolyte imbalance. If untreated, low concentrations of sodium will result in death during the neonatal period. Approximately 25% of cases of P450c21 deficiency can synthesise sufficient amounts of aldosterone to prevent salt loss and are classified as 'simple virilizers' (Fiet et al 1989). Nonclassical P450c21 deficiency is common (Azziz et al 1989) and is caused by allelic variants of CYP21 that have between 20 and 50% wild type P450c21 enzyme activity. Patients present some time after birth with evidence of androgen excess (Mahesh et al 1968; Speiser et al 1985) or may remain asymptomatic (Cryptic).

Cytochrome P450c21 deficiency is inherited as a monogenic recessive disorder linked to the class III region of the HLA complex on chromosome 6p21.3 next to the C4A and C4B components of serum complement (White et al 1985a). This region includes the functional P450c21 gene, CYP21 and a highly homologous pseudogene, CYP21P, both genes contain 10 exons and 9 introns, 98% and 96% homologous respectively (Higashi et al 1986; White et al 1986). All mutant alleles identified to date have been generated
through intergenic recombination, 20% of which have a 30kb deletion that results in a single non-functional chimeric gene, with the 5' end corresponding to CYP21P and the 3' end to CYP21 (White et al 1985b). 75% of alleles have apparently been generated by gene conversion events, where CYP21 carries one or more deleterious mutations normally found in the CYP21P pseudogene. Although the mechanisms by which these mutation events are generated is not well understood, it has been suggested that large deletions are generated by unequal meiotic crossing over, whereas gene conversion is associated with mitotic recombination (Collier et al 1993; Tusie-Luna et al 1995). The very high proportion of intergenic crossing over suggests these genes contain one or more hotspots for recombination or the HLA region may be subject to relatively high rates of gene conversion.

Cytochrome P450c11 Deficiency

P450c11 deficiency is the second most common cause of congenital adrenal hyperplasia with an incidence of 1 in 100,000 live births, approximately 5% of adrenal steroid defects (White et al 1987a; White et al 1987b; Zachmann et al 1983). Patients with this defect are unable to convert 11-deoxycortisol to cortisol (Figure 1.6). Elevated ACTH cause accumulation of steroid precursors that may be utilised in adrenal androgen biosynthesis. Patients present with similar symptoms as P450c21 deficiency, and both sexes show signs of mild androgen excess, rapid somatic growth with premature epiphyseal closure resulting in short adult stature. A parallel defect in the mineralocorticoid pathway leads to the accumulation of deoxycorticosterone causing hypertension. This clinical feature distinguishes P450c11 from P450c21 deficiency.

There are two homologous P450c11 enzymes encoded by CYP11B1 and CYP11B2 on chromosome 8q21-22 (Momet et al 1989). CYP11B1 is expressed in the zona fasciculata under ACTH regulation and is required for cortisol biosynthesis. CYP11B2 has three activities required for aldosterone synthesis in the zona glomerulosa and is mainly regulated by the renin-angiotensin system. P4501c11 deficiency may be inherited as an autosomal recessive disease, however there is a form of hypertension that is inherited in an autosomal dominant disorder. It can be distinguished from primary aldosterism by suppression of aldosterone hypersecretion with dexamethasone. This suggests that CYP11B2 may be under appropriate control by ACTH. Unequal intergenic recombination between CYP11B1 and CYP11B2 results in a hybrid gene with the 5' terminal region of CYP11B1 fused to the coding region of CYP11B2 (Lifton et al 1992).
3β-HSD deficiency

3β-HSD converts 3β-hydroxy-Δ5-steroids to their corresponding Δ4-3-oxo steroids (Figure 1.3 and 1.6). Deficiency of the enzyme results in a rare form of CAH and is of interest because the biosynthesis of all classes of steroid hormones in the adrenal cortex and gonads are impaired (Figures 1.6-1.8), characterised by elevated serum ratios of Δ5-ene to Δ4-ene steroids and increased secretion of Δ5 metabolites, pregnentriol and 16α-pregnenetriol in urine. Deficiency in cortisol synthesis causes increased secretion of ACTH and hyperplasia of the adrenal gland, and low androgen levels lead to an increase in luteinising hormone (LH) and follicle stimulating hormone (FSH) synthesis and secretion. As with P450c21 deficiency there are several clinical phenotypes. Neonates affected with severe classic CAH exhibit signs of electrolyte imbalance to varying degrees which may be fatal if not diagnosed and treated early (Bongiovanni et al 1962). Male infants present with incomplete masculinisation of the external genitalia, due to low levels of testosterone, while affected females show near normal sexual differentiation or slight virilization (Rheaume et al 1992; Chang et al 1993). In contrast non-salt losing 3β-HSD deficiency is normally diagnosed sometime after birth, males present with ambiguous genitalia or hypospadias and females present with mild androgen excess but (both sexes) synthesise sufficient aldosterone to avoid salt wasting crises. By analogy with P450c21 deficiency it has been suggested that non-classical 3β-HSD (NC-3β-HSD) deficiency is an allelic variant of classical 3β-HSD deficiency (Pang et al 1985) that was initially estimated to be more frequent than NC-21 OH (Zerah et al 1991) on the basis of elevated serum levels of basal and post ACTH stimulation of Δ5-ene steroids or by elevated urinary 5-ene steroid metabolites. No 3β-HSD mutation has been found in patients diagnosed with NC-3β-HSD deficiency suggests this disorder does not result from mutations in 3β-HSD type II enzyme (Mathieson et al 1992; Zerah et al 1994).

3β-HSD deficiency is inherited as an autosomal recessive disorder. As discussed more fully in section 1.9, in humans at least two homologous, closely linked genes are expressed, HSD3B1 and HSD3B2. These genes encode the 3β-HSD type 1 enzyme expressed principally in placenta and skin, and the type 2 enzyme in adrenals and gonads (The et al 1989, Rheume et al 1991). Consistent with the biochemical profiles, mutations have been detected in various HSD3B2 genes from patients, showing that 3β-HSD type II activity expressed predominantly in adrenals and gonads is affected Table 1.1 (Rheume et al 1992; Rheaume et al 1994, Simard et al 1993b). The complete nucleotide sequence of HSD3B1 has been determined from patients with 3β-HSD deficiency, no deleterious mutations have been identified suggesting that 3β-HSD type I is functional in these patients. The increased levels of serum Δ4-steroids and their metabolites in patients with classic 3β-HSD deficiency are explained by peripheral (extra-adrenal and extra-
gonadal) 3β-HSD activity catalysing elevated levels of Δ^5 substrates resulting from 3β-HSD type II deficiency in adrenals and gonads. Whether 3β-HSD type I is solely responsible for peripheral activity has yet to be determined. It has been known for some time from Southern blot evidence that a number of homologous 3β-HSD gene sequences other than HSD3B1 and HSD3B2 exist in the human genome. It is not known whether these sequences represent the human 3β-HSD ketosteroid reductase activity, peripheral 3β-HSD activity or pseudogenes. Furthermore recombination, whether intergenic or apparent gene conversion plays a major role in the generation of mutations in cytochrome P450c21 and P450c11 deficiency. The extent to which the as yet unidentified 3β-HSD sequences are involved in the generation of 3β-HSD type II mutations is unknown.

Most of the defects in 3β-HSD deficiency are due to nucleotide changes that cause single amino acid substitutions or premature termination codons within the open reading frame encoding 3β-HSD type II (Table 1.1). Less frequently frameshift mutations have been identified. The majority of these mutations are localised to residues highly conserved throughout the vertebrate 3β-HSD isozyme family and result in partial or complete impairment of 3β-HSD type II enzyme activity.
Table 1.1. Catalogue of 3ß-HSD type II mutations.

<table>
<thead>
<tr>
<th>3ß-HSD type II amino acid mutation</th>
<th>3ß-HSD type II DNA mutation</th>
<th>Phenotype (% activity of mutant enzyme)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G15D Rheaume 1995</td>
<td>GGT -&gt; GAT (Missense)</td>
<td>Salt wasting male pseudohermaphrodite (None detectable)</td>
</tr>
<tr>
<td>A82T (Vas) Medonca 1994</td>
<td>GCC -&gt; ACC (Missense)</td>
<td>Nonsalting male, female pseudo; premature pubarche; cryptic (None detectable)</td>
</tr>
<tr>
<td>N100S (Sta) Mebarki 1995</td>
<td>AAT -&gt; AGT (Missense)</td>
<td>Non salting salt wasting male, female pseudo</td>
</tr>
<tr>
<td>Splicing mutation</td>
<td>G -&gt; A (nt 6651)</td>
<td>Male pseudo without salt wasting; Premature pubarche</td>
</tr>
<tr>
<td>L108W Sanchez 1994</td>
<td>TIG -&gt; TGG (Missense)</td>
<td>Salt wasting male pseudohermaphrodite</td>
</tr>
<tr>
<td>G129R Rheaume 1994</td>
<td>GGG -&gt; AGG (Missense)</td>
<td>Nonsalting male pseudo; Premature pubarche</td>
</tr>
<tr>
<td>E142K Simard 1993</td>
<td>GAA -&gt; AAA (Missense)</td>
<td>Male, female pseudo; varying salting loss (None detectable)</td>
</tr>
<tr>
<td>W171X Rheaume 1992</td>
<td>TGG -&gt; TAG (Non-sense)</td>
<td>Male hypospadias and gynecomastia; Female salting wasting (None detectable)</td>
</tr>
<tr>
<td>P186L Sanchez 1994</td>
<td>CCC -&gt; CTC (Missense)</td>
<td>Salt wasting male pseudohermaphrodite</td>
</tr>
<tr>
<td>186/C/187 Rheaume 1992</td>
<td>(Frameshift)</td>
<td>Male hypospadias and gynecomastia; Female salting wasting</td>
</tr>
<tr>
<td>L205P*</td>
<td>CTG -&gt; CCG (Missense)</td>
<td>Female pseudo; premature pubarche</td>
</tr>
<tr>
<td>S213G</td>
<td>AGT -&gt; GGT (Missense)</td>
<td>Female pseudo; premature pubarche</td>
</tr>
<tr>
<td>K216E</td>
<td>AAG -&gt; GAG (Missense)</td>
<td>Female pseudo; premature pubarche</td>
</tr>
<tr>
<td>A245P Simard 1993</td>
<td>GCC -&gt; CCC (Missense)</td>
<td>Nonsalting (11%)</td>
</tr>
<tr>
<td>V248N Chang 1993</td>
<td>GTC -&gt; AAC (Missense)</td>
<td>Salting male, female pseudohermaphrodite (None detectable)</td>
</tr>
<tr>
<td>R249X Chang 1993</td>
<td>CGA -&gt; TAG (Non-sense)</td>
<td>Salting male, female pseudohermaphrodite (None detectable)</td>
</tr>
<tr>
<td>Y253N Simard 1993</td>
<td>TAT -&gt; AAT (Missense)</td>
<td>Male, female pseudo; varying salting loss (None detectable)</td>
</tr>
<tr>
<td>Y254D Sanchez 1994</td>
<td>TAC -&gt; GAC (Missense)</td>
<td>Non salting female pseudo. (Only one allele isolated) (None detectable)</td>
</tr>
<tr>
<td>T259M (De Porto)</td>
<td>ACG -&gt; ATG (Missense)</td>
<td>Salting female pseudo</td>
</tr>
<tr>
<td>273AAA Simard 1994</td>
<td>(Frameshift)</td>
<td>Salting female pseudo</td>
</tr>
<tr>
<td>315AC Zhang 1996</td>
<td>(Frameshift)</td>
<td>Salting female pseudo</td>
</tr>
</tbody>
</table>
1.9 Molecular biology of 3β-hydroxysteroid dehydrogenase Δ⁵/Δ⁴ isomerase (3β-HSD).

Following 3β-HSD purification from human placenta (Thomas et al 1988) and development of polyclonal and monoclonal antibodies, cDNAs from different vertebrate species have been isolated and characterised by a number of groups. Multigene families for 3β-HSD have been described in myomorph rodents, which encode enzymes with distinct functional properties. 3β-HSD type I-III in mouse and I, III and IV in rat have dehydrogenase/isomerase activities and utilise NAD⁺ as cofactor (Bain et al 1991; Zhao et al 1991; DeLaunoit et al 1992; Keeney et al 1993). 3β-HSD types IV and V in mouse and type III in rat have an exclusive keto steroid reductase (KSR) activity, which is NADP-dependent and is most likely associated with the inactivation of reduced hormone derivatives such as 5α-dihydrotestosterone (Clarke et al 1993a; Clarke et al 1993b; Simard et al 1993a; Abbaszade et al 1995; see sections 1.10 and 6.1).

In humans two highly homologous 3β-HSD cDNAs have been cloned and map to the chromosomal region 1p13.1, syntenic with the mouse 3β-HSD gene family on chromosome 3 (Bain et al 1993). RNase protection assays have been used to estimate the abundance and specific location of 3β-HSD transcripts within human tissue (Rheaume et al 1991). Human 3β-HSD type I mRNA is principally expressed in placenta and skin and at low levels in liver, breast and testis, while 3β-HSD type II mRNA is found almost exclusively in the adrenal cortex, testis and ovary (Labrie 1992). The two cDNA sequences show 93.5% identity within the coding region. 3β-HSD type I cDNA encodes a protein of 373 amino acids and 3β-HSD type II encodes a protein of 372 amino acids. Type I and type II cDNAs expressed transiently in non-steroidogenic COS-1 cells convert multiple Δ⁵-3β-hydroxysteroids substrate to their corresponding Δ⁴-3-ketosteroids proving that both the dehydrogenase and isomerase activities reside within a single protein (Lorence et al 1990, Rheume et al 1991). Both enzymes have ketosteroid reductase activity, and no other ketosteroid reductase enzyme has been described.

Corresponding genes, HSD3B1 and HSD3B2 were subsequently identified from human leukocyte genomic libraries using 3β-HSD type I cDNA as a probe (Lorence et al 1991; Lachance 1991). As illustrated in figure 1.9, the two human genomic 3β-HSD genes are similar containing 4 exons and 3 introns, with conserved intron-exon splice junctions. The first exon contains 53 bp of the 5' non-coding, while exon 2 contains 85 bp of the non-coding region as well as the nucleotide sequence of the first 48 amino acids of 3β-HSD type I. Exon 3 is relatively small encoding 55 amino acids. Exon 4 contains nucleotides encoding the remaining 270 amino acids as well as approximately 400 bp of the 3'
untranslated region. HSD3B2 is similar except that the encoded protein utilises a different initiator methionine; therefore 3β-HSD type II is one amino acid smaller.

**HSD3B1 (≈ 7.7 kb fragment)**

![Diagram of HSD3B1](image)

**HSD2B2 (≈ 7.7 kb fragment)**

![Diagram of HSD2B2](image)

Figure 1.9. Intron/exon structure and partial restriction map of the 3β-HSD genes (not to scale). Exons are represented by boxes numbered I, II, III and IV. Open boxes represent 5' and 3' untranslated sequences, Introns are labelled A, B and C. Common restriction sites, and the translational initiator and stop codons are indicated.

The HSD3B1 gene was assigned by fluorescent in situ hybridisation (FISH) to human chromosome region 1p13.1 (Berube et al 1989; Morrison et al 1991), and HSD3B1 and HSD3B2 were shown to be in close genetic linkage (Russell et al 1994; Lorence et al 1991). Several bands of hybridisation were detected when Southern blots of human genomic DNA were probed with 3β-HSD type I exon specific probes (Lachance et al 1990; Lachance et al 1991; Russell 1993) suggesting that there are other 3β-HSD-like sequences in the human genome. Genomic Not I and Sac II restriction digests resolved on pulse field gels were hybridised under low stringency conditions with 3β-HSD type I and II cDNAs. A single band of hybridisation of about 290 kb was detected after digestion with Sac II restriction enzyme (Morrisette 1993). This suggests that the human 3β-HSD gene family is located within this 290 kb Sac II DNA fragment, consistent with the mouse chromosomesal localisation data (Bain et al 1993).

1.10 Structure-function predictions of 3β-HSD isozymes.

The 3β-dehydrogenase and Δ5/Δ4 isomerase reactions are encoded by separate genes in *Pseudomonas* bacteria (Yin et al 1991), however, in vertebrate tissues the two activities are catalysed by a single protein (Lorence et al 1990, Rheume et al 1991). The evolutionary significance of constraining these activities within a single protein is unknown, however mammalian 3β-HSD isozymes are highly homologous, showing less
variation across species than other steroidogenic gene families show within a single species (Krozowski 1992).

A number of 3β-HSD cDNAs have been identified from mammalian species. The amino acid sequences of mammalian 3β-HSD isozymes reveal little significant sequence similarity with 11β- and 17β-hydroxysteroid dehydrogenases (HSD), cytochrome P450 enzymes or steroid-binding proteins. However, sequence similarity is found between mammalian 3β-HSD isozymes and barley A1 protein, and a bacterial (Norcardia) cholesterol dehydrogenase as well as open reading frames in vaccinia virus (Baker & Blasco 1992).

Based on alignments of predicted amino acids of mouse, rat, human, macaque and bovine 3β-HSD cDNAs, Abbaszade et al (1995) suggested that several gene duplications have occurred during the evolution of the 3β-HSD gene family in these species. One such event arose during the evolution of primates, an ancestral 3β-HSD gene was duplicated specifically in the primate lineage to yield the two human 3β-HSD genes identified to date, HSD3B1 and HSD3B2. Another 3β-HSD gene duplication event occurred during the evolution of rodents before the divergence of the mouse and rat ancestor and several gene duplications took place independently after divergence in these lineages. Considering the presumed evolutionary constraints necessary for 3β-HSD activity, it seems likely that amino acids conserved across mammalian species are important for enzyme activity.

Examination of the mammalian 3β-HSD isozymes shows a typical βαβ dinucleotide binding fold in the amino terminal region (residues 2-36) (Wirenga et al 1986). Affinity alkylation of purified bovine adrenal 3β-HSD with the cofactor analogue 5'-[p-(fluorosulphonyl) benzoyl][adenosine-814C] confirms the amino terminal region for cofactor (NAD+) binding, and that residues 20 and 278 are in close proximity in the 3-dimensional structure of the protein (Rutherfurd et al 1991a; Rutherfurd et al 1991b). Residue 36, aspartic acid (D-36), is the acidic amino acid predicted to hydrogen bond with all known NAD-dependent dehydrogenases, and is identified in all vertebrate 3β-HSD isozymes except mouse IV, V and rat III. Mouse IV, V and rat III enzymes are characterised by either a phenylalanine (F) or tyrosine (Y) at residue 36, and do not display oxidative activity on the classical substrates pregnenolone and dehydroepiandrosterone (DHEA) (Figure 1.3 and 1.4). Using NADPH as cofactor, the preferred substrates are 5α-dihydrotestosterone and other 3-keto saturated steroids. These enzymes are classified as ketosteroid reductases (KSR).
Using site directed mutagenesis a point mutation was introduced in the predicted βαβ dinucleotide binding region of the rat 3β-HSD type III isoform replacing tyrosine (Y) with aspartic acid (D) at amino acid position 36 (Turgeon 1995). Transient expression of rat III Y36D in tissue culture cells shows a decrease in activity with NADPH as cofactor, but a marked improvement of catalytic activity utilising NADH as cofactor. Rat III Y36D shows a similar affinity for NADH to rat 3β-HSD type I, a classic dehydrogenase/isomerase for cofactor. This suggests that residue 36 may be diagnostic of cofactor specificity, D-36 characterises the enzyme utilising NADH as cofactor and Y-36 (or F-36 in mouse) utilising NADPH as cofactor.

Affinity radioalkylation experiments of human 3β-HSD type I protein has identified peptides and amino acids involved in substrate and cofactor binding. Both substrate (pregnenolone Thomas et al 1993) and cofactor (NADH Thomas et al 1994) protect the same internal tryptic, residues 251-274 and 176-186. Protection and radioalkylation of residues His 262 and Cys 183 suggest that they are involved in substrate and cofactor binding, and are in close proximity in the 3 dimensional structure of the protein, forming a bifunctional catalytic site. Based on these results, Thomas 1995 proposed that initially the 3β-HSD enzyme is in the dehydrogenase configuration as it oxidises 3β-hydroxy-5-ene steroids to their 3-oxo-5-ene steroid intermediates and NAD⁺ is reduced to NADH. End product NADH induces a conformational change that activates isomerase activity to produce the 3-oxo-4-ene steroid. Once the reaction is complete, product steroid and NADH dissociate, and the enzyme returns to the dehydrogenase conformation (Thomas et al 1995).

1.11 Tissue distribution of 3β-HSD activity in humans.

The tissue distribution of 3β-HSD protein has been investigated using polyclonal and monoclonal antibodies raised against purified human placental 3β-HSD protein (Mueller et al 1987, Thomas et al 1988). These antibodies have been used in immunohistochemical studies of a number of species including human, rodent and bovine (Sasano et al 1990; Naville et al 1991; Riley et al 1992; Dumont et al 1992).

In humans, high levels of immunoreactivity to 3β-HSD was observed in placental tissues, specifically syncytiotrophoblast cells, extravillous cytotrophoblasts and cells of the chorion laeve of normal term placenta (Riley et al 1992). A description of 3β-HSD protein localisation in the placenta is also described in section 1.13. Immunoreactive protein is observed in fetal testis at an early stage of development (Pelletier et al 1992) consistent
with its critical role in the biosynthesis of testosterone. In the female fetus ovarian 3β-HSD immunoreactivity appears around the fourth month during the period of genital and urethral separation (Pelletier et al 1992).

Adult interstitial Leydig cells of the human testis contain significant amounts of immunoreactive protein. In small developing follicles of the ovary, immunoreactivity was observed only in the theca interna, but staining became apparent in granulosa cells as the menstrual cycle progressed (Sasano et al 1990). In the adult adrenal cortex all three morphological zones, zona glomerulosa, zona fasciculata, zona reticularis exhibited immunoreactivity (Sasano et al 1990) as well as a few peripheral tissues such as sebaceous glands in skin (Dumont et al 1992).

3β-HSD activity is also associated with many other peripheral tissues including kidney, lung and brain (Milewhich et al 1991). However, protein was not readily detected by immunoblot analysis, and the identity of these 3β-HSD isozymes remains to be established.

1.12 Subcellular localisation of 3β-HSD.

3β-HSD enzyme activity localises both to the endoplasmic reticulum, with the active sites for substrate and cofactor binding facing the cystol (Alvarez et al 1994), and to the inner mitochondrial membrane where the active sites face the matrix space (Sauer et al). 60% of total 3β-HSD activity localised to microsomes, and 40% to mitochondria. 3β-HSD does not contain a mitochondrial localisation signal, although other steroidogenic enzymes localising to mitochondria e.g. P450scc and P450c11 contain an amino-terminal peptide leader sequence. Purified 3β-HSD enzyme from both microsomal and mitochondrial compartments exhibit identical biophysical properties, suggesting that they were the same protein localised to different compartments. In addition the amino-terminal amino acid sequences determined for enzyme purified from each organelle were identical (Thomas 1989). Using bovine adrenal cortical cells, Cherradi et al 1993 supported the dual localisation of a single 3β-HSD protein to the endoplasmic reticulum and mitochondria. However, substrate specificities of 3β-HSD differed between the two organelles, suggesting that 3β-HSD activity could be modulated in the bovine adrenal cortex, thereby affecting substrate utilisation (Charradi et al 1994).
1.13 Background and objectives of project.

At the interface between the developing conceptus and the maternal circulation, the fetal cells surface membrane may well play a role in providing immunological protection for the genetically alien fetus. In the human placenta, cells of the villous syncytiotrophoblast and non-villous cytotrophoblast are in immediate contact with maternal blood (Boyd & Hamilton 1970). Much interest has been shown in identifying proteins expressed by the surface membrane of fetal cells since they may provide an understanding of the mechanisms which allow the fetus to survive. Mouse monoclonal antibodies were raised to human first trimester trophoblast cells, including FDO 161G. which bound to villous syncytiotrophoblast and non-villous trophoblast in first trimester and term placentae (Mueller et al 1987). However, after examining reactivities with a panel of other cells and tissues, it was shown that FDO 161G was not trophoblast specific. A further mouse monoclonal antibody, FDO 26G, raised against protein which had been purified from placental villous tissue, had more restrictive tissue reactivity. Immunocytochemically, it reacted with placental villous trophoblast and cells of the adult adrenal cortex, but reacted weakly or not at all with placental extravillous trophoblast and adult testicular Leydig interstitial cells (Hawes et al 1994).

In chapter three the main objectives were to identify a cDNA encoding the FDO 161G reactive protein and confirm reactivity to FDO 26G. As part of a collaborative effort, a cDNA was identified and sequenced and identified as 3β-hydroxysteroid dehydrogenase Δ5/Δ4 isomerase type I (3β-HSD type I). Furthermore, FDO 26G reactivity was confirmed to this protein in Chinese hamster ovaries (CHO) cells.

As indicated in sections 1.9, two 3β-HSD enzymes have been described in humans; 3β-HSD type I is principally expressed in placenta and skin, while 3βHSD type II is expressed in adrenals and gonads (Lachance et al 1990, Rheume et al 1991). The two forms are 95% similar in primary amino acid sequence and antibodies raised against human placental 3β-HSD react immunocytochemically with 3β-HSD in all four steroidogenic tissues. Such tissue specificity has been demonstrated by the monoclonal antibody (mAb) FDO 161G (Mueller et al 1987, Nickson et al 1991). However, the staining patterns of FDO 26G were unexpected, providing evidence for expression of a further 3β-HSD isozyme in certain placental tissues and in adult testis. From these experimental observations, two hypothesis for the extinction of FDO 26G reactivity in placental tissues seem likely; either a post-translational modification within the region of monoclonal antibody binding (the epitope) or the expression of as yet unidentified 3β-HSD gene that encodes a different amino acid sequence over the FDO 26G epitope.
Several bands of hybridisation are detected when Southern blots of human genomic DNA are hybridised with 3β-HSD type I cDNA. The intensity and pattern of hybridization suggest that a number of 3β-HSD genes other than HSD3B1 and HSD3B2 are present in the human genome (Russell 1993). Furthermore when screening PCR amplified fragments of patient DNA for mutations in the genes encoding 3β-HSD types I and II, it was relatively common to amplify novel but closely related 3β-HSD exon sequence (Russell 1993). This is consistent with the identification of multigene families for 3β-HSD in rodents, expressed sex-specifically in peripheral and steroidogenic tissues (Bain et al 1991; Zhao HF et al 1991; DeLaunoit et al 1992; Keeney et al 1993; Clarke et al 1993b; Simard et al 1993a; Abbaszade et al 1995).

Two experimental strategies were devised to address the central questions of the extinction of FDO 26G reactivity in certain placental tissue and the number of human 3β-HSD genes in the family.

Western immunoblot analysis implicated the FDO 26G epitope as the site of modification (Hawes et al 1994). In chapter four, the main objective was to map the FDO 26G epitope on the 3β-HSD type I protein using lacZ gene fusions and synthetic peptides. Identification of the epitope may suggest a post-translational modification consensus sequence, which could be tested using modified peptides.

The size of the human 3β-HSD gene family would be investigated by homology probing with 3β-HSD type I cDNA under nonstringent conditions. In chapter 5, a number of positive clones were identified from two genomic libraries. The main objective in this chapter was to identify a corresponding genomic clone of a previously amplified novel putative exon 3 fragment from a patient genomic library and a normal control genomic library, and to investigate whether this novel gene was responsible for idiopathic hirsutism in family Phe (Russell 1993). The remaining genomic clones would be investigated in chapter 6. These clones were screened for novel members of the gene family by PCR amplification followed by denaturing gradient gel electrophoresis (DGGE) analysis. Various combinations of PCR primers were utilised, particularly those that previously amplified other novel PCR products from genomic DNA (Russell 1993). It was hoped that genomic clones could be identified that included these fragments.

DGGE is one of the most powerful and simple methods for detecting mutations in DNA without sequencing. The addition of GC clamps to the end of one of the primers (Sheffield et al 1989) and the additional refinement of screening for heteroduplex DNA molecules provides a detection system close to 100% efficiency (Mendonca et al 1984; Lerman & Silverstein 1987). It therefore has adequate power to distinguish between
different members of a multigene family which can be amplified with same PCR primer pair. Fortunately the DGGE conditions had previously been optimised for these sequences (Russel 1993). However, with the onset of technology, as sequencing directly from PCR products became more simple without the additional requirement for cloning, this method would also be fully exploited in identifying novel members.
Chapter 2
Materials and Methods

2.1 Random priming of DNA: 20ng of DNA was labelled with $[^{32}\text{P}]\text{dCTP}$ (3000Ci/mm; Amersham) by the random primer method, using a Ready-To-Go DNA labelling kit (Pharmacia), to a specific activity of $10^9\text{dpm/ug}$. The reaction mixture (final volume 50ul) consisted of 20ng DNA in 45ul water, 5ul $[^{32}\text{P}]\text{dCTP}$ (10uCi/ul), and Reaction Mix (dATP, dGTP, dTTP, FPLC pure R Klenow fragment 4-8 units), and random oligodeoxyribonucleotides, primarily 9-mers). DNA was heated at 100°C for 2 minutes before adding to the above mix. The reaction was incubated at room temperature overnight then the reaction was stopped using 2ul of 0.5M EDTA. After adjusting the volume to 100ul with TE the labelled DNA was separated from the unincorporated $[^{32}\text{P}]\text{dCTP}$ by chromatography on a Sephadex G-50 column, using a 2cm x 1cm bed ( height x diameter ).

2.2 PCR primer sequences: A comprehensive schematic representation of PCR/sequencing primer hybridisation sites is presented in figure 6.1.

38-HSD type I (Numbers represent coordinates for 38-HSD type I genomic sequence)
Exon 1/2
M- 1354- AGTGACATAAAGCTTCAG -1370
L- 1545- TAAACCATTGACATCTC -1561
I- 1585- CCTAGAATCAGATCTGC -1601
Y- 1654- ATGACGGGCTGGAGCTGCCTTGAC -1679
K- 1791- TTCCTCTCTCAATTCTGGTCCGAAGGCC -1764
J- 1855- CCACATACATGCAGTGT -1839

Exon 3
A- 5655- CCAATGACCTGACCTGT -5671
P- 5662- CCTGACCTGTGTTCTTG -5679
D- 5682- AACTCCAGAACAAGACC -5698
Q- 5786- AGGCCGGTGCGGATGACT -5770
B- 5881- CCACCTTGCTGCATCTC -5865

Exon 4
E- 7955- CGTGGTGGCACCTCTT -7971
V- 8031- GAGGCCTGTGCAAGGC -8047
3β-HSD type II (Numbers represent coordinates for 3β-HSD type II genomic sequence)

Exon 1/2
T - 1219 - AGAGCATAAAGCTCCAG -1235
S - 1855 - CCACATACATGCAGTGT -1839

Exon 3
A - 5655 - CCAATGACCTGACCTGT -5671
C - 5714 - CCAACTTGTTTTATCTC -5698

Exon 4
F - 7849 - TGTGGTTGCGCTCCCTT -7865
Z - 8270 - AGACTGGGTTGACTGAT -8254

Unless otherwise indicated, the non-stringent PCR reactions conditions were as follows:

94°C 90 sec
45°C 30 sec 35 cycles.
72°C 60 sec

2.3 Screening of human placental expression library λ gt 11 and PCR reaction conditions: A lambda gt11 expression library containing 1.5x10^6 independent clones was screened using a partial cDNA fragment labelled by random priming. Hybridisation was carried out overnight at 65°C in 5x SSC, 5x Denhardt's solution, 0.2% SDS and 100μg/ml salmon sperm DNA, followed by a single wash at room temperature in 1x SSC, 0.1% SDS for 15 minutes and 2 washes at 65°C in 1x SSC, 0.1% SDS each for 15 minutes. Selected plaques were picked as BBL- agarose plugs (Maniatis et al 1982) which were twice freeze thawed in 50μl to release bacteriophage DNA. The supernatant was used as a source of template for PCR reactions (94oC 30 sec, 45oC 30 sec, 72oC 30 sec, 35 cycles), which were carried out in 50μl final volume in Taq buffer (50mM-KCL/ 100mM Tris-HCL pH8.4/ 15mM-MgCl2/ O.1% gelatin), containing 8μl of a dNTP solution (1.25mM dATP/ 1.25mM
dCTP/ 1.25mM dGTP/ 1.25mM dTTP) and 1µl of PCR primers that hybridise to the lambda arms on either side of the EcoRI cloning site. 1 unit of Taq was used/reaction.

2.4 Agarose gel electrophoresis: Unless otherwise indicated, all agarose gels were run in TBE buffer (89mM Tris, 89mM boric acid, 2mM EDTA). Ethidium bromide was added to each gel at a concentration of 0.5µg/ml. The sample loading buffer used consisted of 0.25% bromophenol blue, 40% (w/v) sucrose in water (6x buffer).

2.5 Mammalian cell culture: The cell lines used in this study were COS-7 and Chinese hamster ovary (CHO.A2H) cells. For transfections and general passaging of cells Glasgow Modified Eagles Medium (GMEM) was used. It was made from a 10x stock and the following supplements were added to 400mls of 1x media.

| 300.0 mls | ddH₂O  |
| 40.0 mls | 10x GMEM |
| 13.4 mls | Sodium bicarbonate (7.5% w/v) stored at 4°C. |
| 4.0 mls | sodium pyruvate (10mM). |
| 4.0 mls | non-essential amino acids (100x liquid) stored at 4°C (see below) |

4.0 mls Penicillin/ streptomycin (1000units/ 10000µg/ ml)
4.0 mls L-Glutamine (200mM) ( for COS-7 media)
Proline ( for CHO.2AH media)
40.0mls Fetal calf serum (FCS) heat treated for 1 hour at 65°C.

Non-essential amino acids (100x liquid)

| L-alanine  | 0.89 |
| L-Asparagine | 1.50 |
| L-Glutamic acid | 1.33 |
| Glycine    | 0.75 |
| L-Proline  | 1.15 |
| L-Serine   | 1.05 |

The media was filter sterilised to remove large protein and lipid precipitates, prior to use.
All cells were grown as monolayers in plastic tissue culture flasks (base area 25cm²) under 10 mls of GMEM at 37 °C in an atmosphere of 5% carbon dioxide. When confluence was reached, cells were passaged as follows. The medium was aspirated and the cells washed with 10mls of PBS (5mM Na₂HPO₄, 7.5mM KH₂PO₄, 2.75 mM KCl, 170mM NaCl) and detached with Trypsin/EDTA (Life Technologies) for 1 minute at 37°C. Fresh cultures were initiated with several dilutions (1/10, 1/20 and 1/50).

2.6 Transfection of mammalian cells using CaPO₄: This method is described by Davis et al 1986. DNA is presented as a calcium phosphate DNA co-precipitate to the cell. Cells were plated 24 hours before transfection, at approximately 3x10⁵ cells / dish. The coprecipitate was formed by mixing plasmid DNA (10μg), 125mM CaCl₂, and 1x HBS (2x HBS 1.5mM Na₂HPO₄, 10mM KCl, 280mM NaCl, 12mM glucose, 50mM HEPES pH7.0) in a final volume of 1ml. The mixture was left at room temperature for 30 minutes, to allow precipitate to form. The precipitate was mixed by pipetting up and down and slowly added to the cells. After 48 hours the cells were trypsinised and either harvested (in the case of transient transfection) or placed under G418 selection to obtain stable transfectants.

2.7 Selection of Stable CHO cell lines: The mechanism of selection is based upon the cells sensitivity to the aminoglycoside antibiotic G418. The resistance gene is located on the pcDNAI/Neo (Clonetech) plasmid. Integration of the plasmid, with the desired insert, into the CHO cell genome allows selection with G418. Cells were incubated in the presence of G418 (800μg/ml). After 10 days of selection, clones were lifted using cloning rings and trypsin, and plated to multi-well plates where they were grown until they reached confluence. They were passed to tissue culture flasks and analysed on western blots (see western immunoblot analysis).

2.8 Immunostaining of transfected cells: Transfected slides were grown in slide flasks. After 24 hours they were washed in PBS, dismantled and fixed in PBS/acetone (3:2) for 5 minutes at 4°C, followed by several washes in PBS to remove any excess acetone and left to dry at room temperature. Slides were placed in a humid slide box and wetted with TBS/BSA. The appropriate dilution of primary antibody (see table 2.1) was added 100μl/slide, followed by an incubation of 30 min at room temperature. Thereafter the slides were washed three times in a circulating bath of TBS. The slides were drained and the second antibody conjugated to FITC, previously diluted in TBS/BSA (see table 2.1), were added (100μl/slide) and incubated at room temperature for 15 minutes, washed and dried as described above. A cover slip was put over the cells with a 50% glycerol/TBS (v/v), 0.1% propidium
iodide solution. The propidium iodide was added to visualise the nucleus. Slides were analysed with a fluorescent microscope, and photographs taken with a standard camera attachment and EKTAR 1000 asa film.

TBS  
9.5g NaCl  
50mls 1MTris-HClpH7.5  
to 1litre with dH2O

TBS/ BSA
1x TBS  
3mg/ml BSA

Table 2.1 Primary and secondary antibody used in immunostaining tissue culture cells.

<table>
<thead>
<tr>
<th>Primary antibody</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>W6/32 (seralab)</td>
<td>Positive control; anti-HLA-A, -B, -C (Barnstable et al 1978) 1/50</td>
</tr>
<tr>
<td>FDO 161G Mueller et al 1987</td>
<td>1/10</td>
</tr>
<tr>
<td>Secondary antibody</td>
<td></td>
</tr>
<tr>
<td>Goat FITC (seralab)</td>
<td>anti-mouse conjugated polyvalent 1/100</td>
</tr>
</tbody>
</table>

2.9 Fusion proteins: Fusion genes were expressed in DS 941 and protein was recovered as inclusion bodies, according to the methods of Ruther & Hill (1983).

2.10 SDS-polyacrylamide gel electrophoresis and western immunoblot analysis: Inclusion body preparations and tissue culture cells were run on discontinuous SDS-lammeli gels. The resolving gel (10%) was made from a 30% stock of acrylamide (acrylamide: bis 30:0.8) diluted in 0.575M pH8.8, 0.1%SDS, and polymerised by the addition of 0.1% APS and 0.05% of TEMED. A solution of 0.1% SDS was overlaid to exclude oxygen (an inhibitor of polymerisation) and ensure a horizontal interface between the stacking and resolving gel. The stacking gel was (5%) was prepared by diluting the 30% stock (30:0.8) in 0.126M Tris pH6.8/0.1% SDS and polymerised by addition of 0.1% ammonium persulphate and 0.5% TEMED, and poured on top of the resolving gel. Samples were diluted in final sample buffer (1:2) and heated to 100°C for 5 minutes before loading. The gels were run in Tris/Glycine buffer (25mMTris-HCL, 192mM Glycine, 20% (v/v) methanol) at 40mA for between 30-60 minutes. A marker of known molecular weight were loaded onto the gel.
Tissue culture cell preparations, fusion proteins and peptide conjugates were blotted onto Hybond C-super (Amersham International plc) in a blotting buffer or 25mM Tris, 192mM glycine, 20% (v/v) methanol, and transferred at 100mA for 15 hours at 23°C. Transfer membranes were blocked with 6% (w/v) non-fat skimmed milk in TBS/ Tween (50 mM Tris-HCL buffer, 150mM NaCl, 0.05% Tween, pH8.0) for 1 hour. Primary antibody, FDO 26G, was incubated for 1 hour. After washing 2x 15minutes in TBS/ Tween. An alkaline phosphatase conjugated, anti-mouse, second antibody (1/5000) was added for 1 hour and washed 2x 15 minutes TBS/ Tween. Reactivity was detected using DAB and BCIP (Promega) as substrate.

2.11 Peptides: These were synthesised by ALta Bioscience (Birmingham University, Birmingham, UK). Phosphopeptides were synthesised by Peptide and Protein Research (Washington Singer laboratories, Exeter, Devon, UK) and their structures were confirmed analytically by mass spectrometry. All peptides were synthesised with N-terminal cysteines to permit conjugation to carrier protein. Peptides based on an internal sequence of 38-HSD were synthesised with a N-terminal amide group to minimise charge differences when compared with the native polypeptide. For binding experiments, peptides were conjugated to either activated keyhole limpet haemocyanin (KLH), or BSA using methods provided by the suppliers (Piece, Rockford. IL, USA), with a molar ratio of hapten to carrier of approximately 1000:1.

2.12 Peptide Competition experiments: Unphosphorylated peptide 354-366 was conjugated to BSA, transferred to Hybond C-super using a slot blotter and blocked with 50mM tris-HCL, 150mM NaCl, 0.05% tween, pH 8.0. FDO 26G was added at a final dilution of 1:100 to aliquots (1ml) of 1-20 000 nM solutions of either phosphorylated or unphosphorylated peptide. A strip of Hybond C-super bearing the BSA-peptide conjugate was immersed in each aliquot for 1 hour, washed, and stained for antibody (see methods for western immunoblot).

2.13 Screening of human genomic libraries in λ gem11 and PCR reaction conditions: Two human genomic libraries were created in λ gem11 (AMS Biotechnology, Witney, Oxon, UK). One was from patient IV-2, the other was from an unrelated normal control. The libraries were screened with 32P-labelled cDNA of 38-HSD type I clone B/3 (Nickson et al. 1991) under non-stringent conditions (hybridization 17h in 5x SSC, 5x Denhardt's solution, 0.2% SDS, 100mg/ml salmon sperm DNA at 65°C; washed twice in 2x SSC, 0.1% SDS at 65°C). From a total of 1.4 x 10^6 plaques (7 genome equivalents), 57 positive plaques hybridised reproducibly through to the secondary stage of screening. Plaques were re-purified through a second round of plating
and hybridisation and then picked into 100µl H2O with a pasteur pipette and boiled to release phage DNA. This DNA was tested as templates for PCR amplification (94°C 30 sec, 45°C 30 sec, 72°C 30 sec, 35 cycles). Amplification products were either cloned in pT7 Blue T-vector (AMS Biotechnology, Witney, Oxon, UK).

2.14 Polymerase chain reaction (PCR): Segments of the genes for 3β-HSD types I and II were amplified from lymphocyte DNA by PCR. PCRs were performed in a Hybaid "Omnigene" for 35 cycles of: [1] 91°C for 30 sec, [2] annealing temperature for 30 sec, [3] 72°C for 30 sec, 35 cycles). The reaction mixture (final volume 50µl) consisted of amplification buffer (50mM KCl, 10mM Tris-HCl (pH9.0), 0.1% Triton X-100, 15mM MgCl₂, 0.2mM of each dNTP, 1uM of each oligonucleotide primer, 0.5-1 unit of Taq DNA polymerase (Promega), and 10-50ng of DNA template.

2.15 Denaturing gradient gel electrophoresis: The DGGE system including materials, reagents, gel preparation, and electrophoresis has been described in detail by Myers et al (1987). Segments of the genes for 3β-HSD types I and II were amplified from bacteriophage DNA by PCR. In each case the sequence of the GC clamp at the 5' end of the PCR primer was that defined by Sheffield et al (1989). DGGE was carried out in 6.5% polyacrylamide gels (37.5:1, acrylamide:bis acrylamide) in a buffer tank containing 91 of TAE buffer (40mM Tris, 20mM sodium acetate, 1mM EDTA, pH 7.4) (Figure 2.1). Denaturing gradient gels (16 x 18 x 0.1 cm) were poured using a Dual Slab Electrophoresis SE 600 gel former (Hoeffer). A Grant Immersion Thermostat (TA 0-150) was used to circulate the electrophoresis buffer at a constant temperature of 60°C. The power pack used was a Kikusui PAB (250-0.25). Appropriate denaturant and acrylamide concentrations were achieved in each gel by the mixing of two stock solutions, 0% denaturant stock (6.5% acrylamide (37.5:1, acrylamide:bis acrylamide) in TAE), and 100% denaturant stock (6.5% acrylamide (37.5:1, acrylamide:bis acrylamide), 7M urea, 40% formamide in TAE). Stock solutions were filtered twice through Whatman No.1 paper and stored in brown bottles at 4°C. 1/200 gel volume of 20% ammonium persulphate and 1/2000 gel volume of TEMED (N, N', N', N'- Tetra methyl ethylenediamine) were used as polymerisation catalysts. DGGE analysis had already been optimised (Russell 1993); Exon 3 DGGE analysis was carried out using 50-70% denaturant gels, while Exon 4B DGGE analysis was carried out with 40-70% denaturant gels. All DGGE gels were run at 80V, 100mA for 20 hours. Best resolution was achieved in the centre and lower two thirds of the gel matrix. DGGE gels were stained by immersion in 280 µg/ul ethidium bromide for 30 minutes.
Samples which exhibited altered mobility on DGGE were either cloned into pT7 vector or directly sequenced.

![Denaturing gradient gel electrophoresis diagram](image)

**Figure 2.1:** Denaturing gradient gel electrophoresis. Schematic diagram to illustrate the equipment used in the running of denaturant gradient gels. The direction of buffer flow is indicated by arrows.

### 2.16 Confirmation of putative clones by PCR

Individual white colonies were selected and grown up in 2ml LB broth (10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl, 950ml H2O) incubated at 37°C overnight. 10ul of culture was added to 40ul water and heated in a boiling water bath for 5 minutes. 10ul of the boiled culture was used as template in the appropriate PCR reaction. PCR products were screened for artefactual PCR mutations by resolution on the appropriate DGGE gel alongside controls (the original PCR product). Only independent clones which resolved indistinguishably from the original genomic template were selected for sequencing. Large scale plasmid preparation of DNA for sequencing was performed using Qiagen columns. Sequencing was carried out using the sequenase version 2.0 DNA sequencing kit (U.S.B.).

### 2.17 pT7 Blue T-vector cloning of PCR products

50ng (0.03pmol) of pT7 Blue T-vector was ligated with 10-50ng (0.04-0.2pmol) of PCR product. PCR products were not band purified before cloning into pT7 Blue T-vector. The reaction mixture (final volume 10ul) consisted of T4 DNA ligase buffer (30mM Tris-HCl (pH 7.8), 10mM MgCl2 10mM DTT), 10mM ATP, 50ng pT7 Blue T-vector, 10-50ng PCR product, and 2-3 Weiss units T4 DNA ligase (B.R.L.). The reaction mixture was
incubated at 15°C overnight. 1 ul of ligation reaction was added to 20 ul of Nova Blue
competent cells (Invitrogen) and placed on ice for 30 minutes. The cells were heated
for 40 seconds in a 42°C water bath and then placed on ice for 2 minutes. 80 ul of SOC
medium (20g bacto-tryptone, 5g bacto-yeast extract, 0.5g NaCl, 20ml of 1M glucose,
950ml H2O) was added and the cells incubated at 37°C for 1 hour. 300 ul more SOC
medium was added and 100 ul portions were spread on LB (Luria-Bertani) agar
plates (10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl, 15g bacto-agar, 950ml
H2O) containing 50 ug/ml ampicillin. For blue/white screening of recombinants 35 ul
of 50 mg/ml X-gal in dimethylformamide and 20 ul 100 mM IPTG in water were spread
on the plates and allowed to soak in for about 30 minutes prior to plating. Plates were
incubated at 37°C overnight. Putative clones were identified by blue/white selection.

2.18 Preparation of DNA: Plasmid DNA was prepared using Wizard mini-prep kits
(Promega).

2.19 Dideoxy DNA sequencing: Double stranded plasmid sequencing was carried
out with a Sequenase kit, USB version 2. Subclones of inserts were sequenced using
primers to both pUC18 and T-vector plasmid and to sections of the insert DNA.

2.20 Direct sequencing of PCR products: PCR products were electrophoresed
through a 1% agarose at 8V/cm gel length for 1-2 hours. Bands were visualised using
long wavelength ultraviolet light and relevant bands were excised using a scalpel.
DNA was purified from the gel slice using a Magic Prep column (Promega).
Sequencing was carried out using the sequenase version 2.0 DNA sequencing kit (U.S.B.).

2.21 Bacteriophage lambda DNA: DNA was prepared using a Lambda DNA
purification kit (Promega, A7290).

2.22 Bidirectional Southern blot analysis of lamda DNA: 1-5 µg of lambda DNA
was restricted with 10 units of restrictron enzyme and the Life Technologies system of
REACT buffers and were carried out at 37°C for 2 hours. DNA gels to be blotted
were run on 1x TBE agarose gels. The gels were depurinated in dilute acid (HCL) to
increase efficiency of transfer and then washed by rocking, in a tray, 2x 15 minutes
denaturating solution (500 mls, 1.5M NaCl, 0.5M NaOH). The gels were then washed
rinsed by rocking, 2x 20 minutes neutralising solution (3.0M NaCl, 0.5 Tris-HCL pH
8.0). A sandwich arrangement was set up as follows. Flat paper towels,
approximately 2-3 inches thick were placed on the bench, followed by 3 sheets of
3MM paper and finally the nylon positively charged membrane (Boehringer
Mannheim). The 3MM paper and the nylon membrane were the same dimensions as the gel to be blotted. The gel was then placed on top followed by nylon, 3MM paper and paper towels. Filters were then baked at 80°C for at least 2 hours.

2.23 Phylogenetic analysis:

Nucleotide multiple alignments were obtained using PILEUP (Program Manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711); trout (S72665), mouse1 (M58567), mouse2 (M77015), mouse3 (M75886), mouse4 (L16919), mouse5 (L41519), rat1 (M38179), rat2 (M38178), rat3 (L17138), rat4 (M67465), macaque (M67467), human type I (M35493), human type II (M77144) and bovine (X17614). If necessary manual adjustments were made using the LINEUP (Program Manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711). Phylogenetic relationship of the new human 3β-HSD sequences relative to other mammalian cDNA sequences was determined using the maximum parsimony method of the PAUP 3.1.1 programme (Swofford 1993).
Chapter 3

Cloning and expression of the cDNA clone B3 (P10), and its identification as 3β-hydroxysteroid Δ5/Δ4 isomerase type I (3β-HSD type I).

3.1 Introduction

The mouse monoclonal antibody designated FDO 161G reacts with a 43kDa protein found in human placental tissues, non-villous trophoblast and villous syncytiotrophoblast (Mueller et al 1987). It has also been shown to react immunohistochemically with cells of the adrenal cortex, sebaceous glands in skin, interstitial Leydig cells of the adult testis and ovarian follicle cumulus cells. However, a further mouse monoclonal antibody FDO 26G raised against the purified human 43kDa protein shows more restrictive staining patterns in trophoblast cell populations and gonads. FDO 26G reacted strongly with villous syncytiotrophoblast, weakly with some trophoblast cells in the chorion laeve and not at all with extravillous trophoblast in cytotrophoblast cell islands. Adult Leydig cells of the testis do not show reactivity with FDO 26G, however, like FDO 161G, FDO 26G reacted strongly with cells of the adrenal cortex. The 43 kDa protein expressed in these tissues attracted interest because of these discrepancies in monoclonal antibody staining (Hawes et al 1994).

Using an affinity-purified sheep anti-161G antibody, a partial cDNA, clone 1/6, had previously been isolated from a λ gt11 placental cDNA expression library (Nickson et al 1991). The main objective of the experiments presented in this chapter were to isolate a cDNA encoding the entire open reading of the 43 kDa protein, and to confirm reactivity to FDO 161G and FDO 26G monoclonal antibodies in tissue culture cells. The staining patterns described above and in section 1.13, could be explained by either the expression of two or more closely related gene products or post-translational modification of a single protein. Therefore a further objective was to estimate the number of different cDNA sequences encoded by positive clones identified in the library screen using DGGE.
3.2. Isolation of B3, a 1700bp placental cDNA clone.

Clone 1/6 insert DNA was radioactively labelled and used as a probe to isolate further clones from the original placental cDNA λgt11 library (1.5x10^6 independent clones). Forty eight positive clones were identified from 150,000 plaques screened, a frequency of 3x10^-4, and partially purified. PCR primers, 695 and 696, were designed to hybridise to regions flanking the polycloning sites of lambda gt11 to allow simple estimation of insert size (Figure 3.1).

Ten positive clones were randomly selected for further analysis, and purified. Positive plaques were picked as BBL-agarose plugs and twice freeze thawed to release bacteriophage DNA suitable for PCR amplification. The approximate size of each PCR product using primers 695, 696 was estimated (Figure 3.2).
Figure 3.2. Estimation of clone insert size by PCR analysis using primers 696 and 695 from ten cDNA clones isolated from the placental cDNA λgt11; P1, P4, P6, P10(B3), P11, P12, P33, P41, P42 and P44 library (approximate sizes are indicated in brackets. Lanes 7 and 8 show the insert size of clone 1/6. P10 (B3) Lane 12 was selected for further analysis. Clone 33 in Lane 5, is not purified to homogeneity. Lane 14 is a negative control for the PCR reaction (no template DNA).

Clone B3 (P10) PCR product (Figure 3.2, Lane 12) was estimated to contain a 1700bp insert and was cloned into pUC18 and sequenced by Debra Nickson (Figure 3.3)(Nickson et al 1991).

Figure 3.3. Partial restriction map and sizes in base pairs of PCR clone B3 (P10) amplified with primers 695 and 696 (1673 bp). Open boxes are untranslated regions, filled boxes are coding sequence. EcoR I sites are maintained from λgt11.

This sequence was used to search Genbank data base and the clone B3 was identified as 3β-hydroxysteroid Δ^5/Δ^4 isomerase type I (3β-HSD type I).
3.3 B3 (3β-HSD type I) PCR cloning into the eukaryotic expression vector pcDNAI/Neo

The purpose of this experiment was to confirm reactivity of FDO 161G to the B3 open reading (ORF) frame isolated from the placental cDNA library. B3 open reading frame was amplified using primer pairs 35/696 introducing a Bcl I restriction enzyme digestion site (Figure 3.4 a). B3 PCR product could then be cloned into the BamHI/EcoRI restriction sites of pcDNAI/Neo (Figure 3.4 b).

![Diagram of PCR amplification and cloning into pcDNAI/Neo](image)

Figure 3.4. Schematic representation of PCR amplification of B3 utilising primers 35/696 and subsequent cloning into pcDNAI/Neo. a) The upstream 5' untranslated and the initial 12 nucleotides encoding the first 4 amino acids of pUCB3 cDNA. Untranslated sequence is presented in lower case. The arrow indicates hybridisation between oligo 35 and upstream sequence, whereas the 'tail' represents extra sequence at the 5' end of the oligo including a Bcl I restriction enzyme site. PCR amplification using oligo 35 and 696 with pUCB3 as template, amplified B3 (3β-HSD type I) open reading frame. b) pcDNAI/Neo is a 7.1 kb eukaryotic expression vector, containing the Neomycin gene. This provides a marker for stable selection in eukaryotic transfections. The Neomycin gene confers Kanamycin resistance in bacteria. Unique polylinker sites are shown in bold. * indicates restriction enzyme sites lost during the cloning of B3 ORF.
Using pUCB3 as template, B3 open reading frame was PCR amplified using primers 35/696 (Figure 3.5 a). pcDNAI/Neo vector was prepared by restriction with enzymes \textit{BamHI} and \textit{EcoRV} (Figure 3.5 b).

![Figure 3.5](image)

Figure 3.5. a) B3 open reading frame was amplified using PCR primers pair 35 and 696.

b) Expression vector pcDNAI/Neo restricted with \textit{BamHI} and \textit{EcoRV}.

B3 ORF cloned into pCDNAI/Neo yields the construct pcDNeoB3. To confirm B3 had been successfully cloned into pcDNAI/Neo, restriction enzyme digestions were compared between pcDNAI/Neo vector DNA and the pcDNeoB3 clone (Figure 3.6).
A number of restriction enzyme site were lost when the B3 open reading frame was cloned into pcDNAI/Neo (Figure 3.4 b). The restriction enzyme digestions in figure 3.6 are diagnostic of B3 open reading frame successfully cloned into pcDNAI/Neo, including the loss of an EcoR I (lanes 5 and 6) and EcoR V (lanes 8 and 9) restriction sites, and the altered restriction digest pattern with BamH I (lanes 11 and 12) and Pst I (lanes 15 and 16).
3.4 Transient expression of pCDNeo/B3 in COS-7 cells and immunostaining with FDO 161G monoclonal antibody.

Transfected COS-7 cells were prepared for analysis using immunostaining and fluorescence microscopy. The integrity of COS-7 cells was determined by immunostaining with anti-HLA class I primary monoclonal antibody W6/32 (Barnstable et al 1978), followed by detection with an FITC conjugated second antibody (Figure 3.7). Propidium iodide stains the nuclei red.

Figure 3.7. Positive control, COS cells immunostained with W6/32 antibody (1/50 dilution).

The construct pcDNeo/B3 was transiently transfected into in COS-7 cells using CaPO₄. After two days, COS-7 transfected cells were fixed and immunostained with FDO 161G, followed by detection with an FITC conjugated second antibody (Figure 3.8)
Figure 3.8. (a) Negative control, COS cells transfected with pcDNAI/Neo vector were immunostained with FDO 161G (1/10 dilution). (b) COS cells transfected with pcDNAo/B3 (10ug), immunostained with FDO 161G (1/10 dilution).
3.5 FDO 26G staining of CHO cells stably transfected with pcDNeo/B3.

pcDNAI/B3 was transfected into Chinese hamster ovary (CHO) cells. Several stable cell lines were selected using resistance to G418 conferred by expression of neomycin (Figure 3.4 b). Immunoreactivity of FDO26G was confirmed by western immunoblot analysis (Figure 3.9).
3.6 Evidence for Bgl II polymorphism from clones isolated in the placental library screen.

As indicated previously, B3 (P10) sequence was used to search Genbank and identified as 3β-hydroxysteroid dehydrogenase Δ^5/Δ^4 isomerase type I (3β-HSD type I). B3 sequence diverged from the published cDNA sequence encoding amino acid residue 338, removing a Bgl II restriction enzyme site. PCR product from B3 and the other clones amplified in figure 3.1, were restricted with Bgl II to determine if the placental library was heterozygous for this restriction enzyme site (Figure 3.10).

![Bgl II analysis of PCR products amplified using primers 696 and 695 from positive clones P1, P4, P6, P10 (B3), P11, P12, P33, P41 and 1/6, clone 44 was not included in this analysis. Six clones contained the Bgl II restriction enzyme site, P1, P4, P6 P11, P12, P41. Four clones did not contain the Bgl II restriction site, 1/6, B3 (P10), P33 and P42. Four clones did not contain the Bgl II restriction site, 1/6, B3 (P10), P33 and P42.

These clones, including 1/6, represent two sequence variants of codon 338 of the open reading frame. There is evidence that this synonomous mutation encoding residue 338 is a common polymorphism (Russell et al 1994). Further sequence analysis by Alan Russell revealed a second base pair missense mutation encoding residue 367, N367T (Russell et al 1994).
3.7 *DGGE* analysis of positive clones isolated from the placental library.

To investigate the remaining clones and confirm they contained 3β-HSD type I sequence, PCR amplification using primers (G, H) was attempted from each of the remaining partially purified clones. PCR primer pair G, H would amplify 480 bp at the 3' end of B3 coding region (Figure 3.11).

![Figure 3.11](image)

Figure 3.11. Structure of the placental cDNA clone B3 (1562bp). The dark box represents coding DNA sequence, the open boxes are 5' and 3' untranslated sequence. Translational start and stop codons are shown along with 33bp of upstream sequence, including the inframe stop and start codons highlighted in bold. Primer pair G and H hybridisation sites are illustrated in relation to two polymorphic sites at the 3' end of the coding sequence. Primer G contained a GC clamp for DGGE analysis.

46 out of 48 clones gave an amplification product of the expected size and these PCR products were analysed using denaturing gradient gel electrophoresis (D.G.G.E.) (Figure 3.12). Haplotype controls (*Bgl* II +, Thr and *Bgl* II -, Asn) had already been identified from the placental cDNA library by restriction digestion analysis (Figure 3.10).
Clone 36 did not amplify with primers G, H. Further PCR amplification was attempted with primers 695 and 696, the PCR product was cloned, sequenced and identified as a 3β-HSD type I clone foreshortened at the 3' end. Clone 38 failed to support amplification with PCR primer pairs (G, H) and (695, 696). An attempt was made to recover this bacteriophage clone by rehybridisation with 1/6 probe, but was unsuccessful. The 3β-HSD type I haplotypes from the positive clones identified in this screen from the placental library are summarised in Table 3.1.
Table 3.1. Summary haplotypes of 3α-HSD type I clones from the placental library amplified with PCR primers (G, H) and analysed on DGGE gels.

<table>
<thead>
<tr>
<th>Placental clone</th>
<th>Approximate size primers 695/696</th>
<th>367 allele</th>
<th>Bgl II + or</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6</td>
<td>1.2 kbp</td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1.6 kbp</td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>1.2 kbp</td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>1.2 kbp</td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>10 (B3)</td>
<td>1.7 kbp</td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>1.6 kbp</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.6 kbp</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>1.2 kbp</td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>33</td>
<td>1.4 kbp</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>36</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>37</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>38</td>
<td></td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>39</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>41</td>
<td>1.6 kbp</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>1.1 kbp</td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>43</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>44</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>45</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>46</td>
<td></td>
<td>Thr</td>
<td>+</td>
</tr>
<tr>
<td>47</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td>Asn</td>
<td>-</td>
</tr>
</tbody>
</table>
Chapter 4

Mapping the monoclonal antibody FDO 26G epitope to the carboxy-terminus of B3 (3β-HSD type I).

4.1 Introduction.

Immunocytochemically FDO 161G reacts with 3β-HSD in trophoblast cells of the placenta, sebaceous glands in skin, cells of the adrenal cortex and gonadal tissue (Mueller et al 1987). However, a further mouse monoclonal antibody FDO 26G raised against purified human placental 3β-HSD shows more restrictive staining patterns in trophoblast cell populations and gonads. FDO 26G reacted strongly with villous syncytiotrophoblast, weakly with some trophoblast cells in the chorion laeve and not at all with extravillous trophoblast in cytotrophoblast cell islands. Leydig cells of the testis do not show reactivity with FDO 26G, however, like FDO 161G, FDO 26G reacted strongly with cells of the adrenal cortex. The monoclonal antibodies FDO 161G and FDO 26G have been shown to react to B3 (3β-HSD type I) protein expressed in COS-7 cells and CHO cells respectively (Figures 3.8 and 3.9).

The differences in staining patterns suggest the expression of a further 3β-HSD isozyme in certain placental trophoblast populations and gonads. RNase protection assays and RT-PCR experiments have shown that 3β-HSD type I transcripts are found in trophoblast cells and 3β-HSD type II is the principal 3β-HSD transcript found in gonads and the adrenal cortex. It seems likely therefore that either a post-translational modification of 3β-HSD type I and type II or the expression of a new gene in the appropriate tissue results in the extinction of FDO 26G reactivity. The main objective in this results chapter is to map the FDO 26G epitope. This would allow either the potential post-translational modifications to be tested, or identify the sequence of amino acids that would be altered in the unreactive tissue, complementing the search for new 3β-HSD genes (results chapters 5 and 6).

Results

4.2 Mapping the FDO 26G epitope using bacterial lacZ fusion genes.

Plasmids pUR290, pUR291 and pUR292 (Ruther U & Mueller-Hill 1983) can be used for expression of genes without their own translational initiation region as fusion proteins with β-galactosidase (lac Z). DNA sequence can be subcloned, inframe, into the carboxy terminus of the lacZ gene using unique restriction sites. The restriction enzyme digestion sites are in different reading frames in different plasmids (Figure 4.1).
Figure 4.1. a) Schematic diagram of pUR290 vector, showing the full length β-glactosidase gene with unique restriction enzyme sites at the 3' terminal region of lacZ. The lacZ sequence is preceded by the lacUV5 promoter allowing regulation of expression of the fusion protein with the lac repressor and IPTG. The ampicillin gene is the selectable marker. b) The unique cloning restriction enzyme sites for the plasmids pUR290, pUR291, pUR292 are shown. The triplet codons represent lacZ reading frame.

Restriction fragments were subcloned from pUCB3 into the bacterial expression vectors, however, suitable restriction sites were not always available and were introduced into the cDNA sequence of B3 by PCR amplification or by site selected PCR mutagenesis.
4.3 Constructs pURMM1 and pURMM2 (Residues 1-241 and 210-373).

B3 was divided into two convenient overlapping 5' and 3' restriction fragments. The 5' fragment is derived from PCR amplification using pUCB3 and 35/696 primers followed by restriction with *Bcl I/Pst I* enzymes (Figure 4.2 a). The 3' fragment is derived from *BamH I/Hind III* restriction of pUCB3 DNA, (Figure 4.2 b). The predicted fusion genes are illustrated in (Figure 4.2 c).

a) Oligo 35

![PCR amplification from the placental clone B3, using PCR primers 35 and 696. Oligo 35 hybridises to sequence upstream of the initiation codon of B3 and introduces a *Bcl I* restriction site at the 5' end of the PCR product. *Bcl I/Pst I* restriction isolates the 5' fragment.](image)

b) Clone B3 subcloned into pUC18, *BamH I/Hind III* restriction isolates the 3' fragment.

![Predicted sequence from the successful cloning of the two restriction fragments pURMM1 and pURMM2. The reading frame of the lacZ gene is indicated by triplet codons. Sequence derived from pUR plasmid is shown in plain text, B3 open reading frame is highlighted in bold. The sequences in lower case are upstream untranslated nucleotides.](image)

pURMM1 (5' construct)
pUR291-TGT CGG Gga tca ctg cta ctt tgg atg gcc ATG ACG GGC TGG AGC-B3

pURMM2 (3' construct)
pUR292-TGT CAG GGG ATG CTG TCA AGT GTT GGA AAG TTC TCC ACT-B3

Figure 4.2. a) PCR amplification from the placental clone B3, using PCR primers 35 and 696. Oligo 35 hybridises to sequence upstream of the initiation codon of B3 and introduces a *Bcl I* restriction site at the 5' end of the PCR product. *Bcl I/Pst I* restriction isolates the 5' fragment. b) Clone B3 subcloned into pUC18, *BamH I/Hind III* restriction isolates the 3' fragment. c) Predicted sequence from the successful cloning of the two restriction fragments pURMM1 and pURMM2. The reading frame of the lacZ gene is indicated by triplet codons. Sequence derived from pUR plasmid is shown in plain text, B3 open reading frame is highlighted in bold. The sequences in lower case are upstream untranslated nucleotides.
The 5' fragment was cloned into pUR291 vector restricted with BamH I/Pst I restriction digestion enzymes. The upstream inframe stop codon is not included (see Figure 4.3). The 3' fragment from pUCB3 was cloned into pUR292 restricted with BamH I/Hind III restriction digestion enzymes (Figure 4.3).

Figure 4.3. Restriction digests of DNA required for pURMM1 and pURMM2 constructs. Lane 1 B3 (35/696) PCR amplified DNA, restricted with Bcl I/Pst I. Lane 2 pUR291 BamH I/Pst I restricted DNA. Lanes 7 and 8 pUCB3 DNA restricted with BamH I/Hind III. Lanes 3 and 4 pUR292 DNA restricted with BamH I/Hind III. Lane 5 λ Pst I marker. Sizes of the restricted pUCB3 DNA are shown.

pURMM1 plasmid transformants that contained the 5' construct were screened using the unique Kpn I restriction digestion enzyme (Figure 4.4 a), while pUMM2 clones were screened using Sst I restriction enzyme (Figure 4.4 b).

Figure 4.4. pURMM1 and pURMM2 transformants. a) pURMM1 transformants were screened using the unique Kpn I restriction digestion site. Lane 14 contains a plasmid that has been linearised. b) pURMM2 constructs were screened using the restriction enzyme Sst I, predicted to excise a 1,900 bp fragment. Lane 2 contains a plasmid that excises a DNA fragment of the correct size.
The plasmids in figure 4.4 a) lane 14 and b) lane 2 were selected for further analysis. The sequence over the polycloning region of each of the constructs was determined to ensure that frame throughout the fusion gene was maintained. The sequence from pUR291 and the pURMM1 construct is presented in figure 4.5.

![Figure 4.5](image)

**Figure 4.5.** a) pUR291 sequence over the polylinker region at the 3' end of the lacZ gene. b) Sequence of construct pURMM1. B3 sequence is highlighted in bold; 5' untranslated DNA is shown in lower case, while coding DNA is presented in capitol letters. The sequence of pURMM1 constructs suggests that the DNA is in frame and will express fusion protein.

The sequence from pUR292 and the pURMM2 construct is shown in figure 4.6. The frameshift encountered by the using the wrong translational vector was corrected by the insertion of an oligonucleotide (Figure 4.6 c).
Figure 4.6. a) Sequence analysis indicates pUR290 was used instead of pUR292. b) The pURMM2 fusion gene contains a frameshift and is one base pair out of frame. c) An oligonucleotide linker DNA sequence was cloned into the pURMM2 fusion gene to correct the frameshift introduced by using pUR290 vector.

pURMM1 and pURMM2 constructs were expressed in *E. coli* and protein was recovered as inclusion bodies, according to the methods of Ruther & Muller-Hill (1983). pURMM1 and 2 inclusion body preparations were tested for reactivity to FDO 26G (Figure 4.13). FDO 26G reacted positively to pURMM2 localising the epitope to residues 241-373 of 3β-HSD type I.

4.4 Construct pURMM3 (Residues 263-373)

The FDO 26G epitope was localised to the 3' region of B3 (3β-HSD type I). As there were few restriction enzyme sites in this region a *Pst* I restriction enzyme site was introduced into the B3 cDNA by site directed mutagenesis by overlap extension using the polymerase chain reaction illustrated in figure 4.7 (a)-(d). The resultant mutant product mB3(1) is shown in figure 4.7 (d). Initially this product was cloned into pUC 18 before being subcloned into pUR292 (Figure 4.8 and 4.9).
Figure 4.7. Site directed mutagenesis by overlap extension using the polymerase chain reaction. a) PCR primers b and c hybridise to 38-HSD type I cDNA, but are synthesised with base mutations encoding amino acid residues 262 and 263, that introduce a PstI restriction enzyme site. b) PCR products 695B and 696C overlap and contain the introduced PstI restriction site. c) These PCR products are band purified from agarose gels, mixed and reamplified using primers 695 and 696. XX represent introduced mutations within oligos b and c.

Introduction of a PstI restriction enzyme site into the cDNA sequence of 38-HSD type I, mB3(1).

b) PCR amplification using primers 695/oligo c and oligo b/696 yield products 695C and 696B.

695C (892bp)      696B (807bp)
Figure 4.8 (a) and (b) show the products of the PCR reaction 695C and 696B, and confirmation of the incorporation of the mutant nucleotides by Pst I restriction digestion (Figure 4.8 (b) lanes 5 and 6). The Pst I/Hind III mB3(1) fragment (Figure 4.8 c) was cloned into pUC18 to give the construct pUCmB3(1) (Figure 4.8 d).

Digestion 6 (Figure 4.8 d) lane 7) excised the required fragment from pUC 18 and was designated pUCmB3(1). This Pst I/Hind III fragment was then cloned into pUR 292 restricted with Pst I/Hind III restriction enzymes to make the construct pURMM3 (Figure 4.9).
pURMM3 constructs were identified by Pst I/Hind III restriction enzyme digestion. (Figure 4.9 c, lanes 2 and 5). The pURMM3 construct was expressed in E.coli and inclusion body preparations (Ruther & Muller-Hill 1983) were tested for reactivity to FDO 26G (Figure 4.13). FDO 26G reacted positively to pURMM3 localising the epitope to residues 263-373 of 3β-HSD type I.

4.5 pURMM4 and 5 constructs. (3β-HSD type I residues 241-285, and 289-373)

Site directed mutagenesis by overlap extension using the polymerase chain reaction was also utilised to construct the translational fusion genes pURMM4 and pURMM5. A premature stop codon (TGA) and a BamH I restriction enzyme site was introduced into the cDNA sequence of B3 (3β-HSD type I) to yield the PCR product mB3(2) (Figure 4.10).
Introduction of a premature stop codon and a BamH I restriction enzyme site into the cDNA sequence of 38-HSD type I, mB3(2).

Figure 4.10. Site directed mutagenesis by overlap extension using the polymerase chain reaction. a) PCR primers b and c hybridise to 38-HSD type I cDNA, but are synthesised with base mutations encoding amino acid residues 286 and 288, that introduce a premature stop codon and a BamH I restriction enzyme site. b) PCR products 695B and 696C overlap and contain the introduced mutation sites. c) These PCR products are band purified from agarose gels, mixed and reamplified using primers 695 and 696. d) The resultant full length PCR product, mB3(2) contains the introduced stop and BamH I restriction site. Incorporation of the mutant nucleotides was confirmed by restriction digest. The sizes of BamH I restriction digests are indicated in (a) and (d).
Figure 4.11 (a) and (b) illustrate the incorporation of mutations into the PCR product to form mB3(2). Restriction enzyme digestion confirms that mB3(2) is cloned into T-vector to make the construct Tvec mB3(2) (Figure 4.11 c).

The T-vector clone containing the mB3(2) fragment was restricted with Pst I/Hind III. The 860 bp band (Figure 4.11 c, lanes 2-5) was purified and cloned into pUR292 to give the construct pURMM4. Tvec mB3(2) was also restricted with BamH I/Hind III (Figure 4.11 c, lanes 7-10). The 720 bp BamH I/Hind III fragment was purified and cloned into
pUR292 to make the construct pURMM5. Transformants were screened using single colony gels (Figure 4.12 a and b). Restriction enzyme digestion confirmed the insert size in pURMM4 and 5 (Figure 4.12 c).

Figure 4.12. Isolation of pURMM4 and pURMM5 constructs. a) Single colony gel of pUR292 plasmid and transformants from the pURMM4 cloning experiment. Lanes 10 and 11 contain plasmids that have different mobilities from the control Tvector plasmid in lane 2. b) Single colony gel of pUR292 plasmid and transformants from the pURMM5 cloning experiment. Lanes 8 and 9 contain plasmids that have different mobilities from the control plasmid in lane 2. c) Restriction digestion of plasmids identified in (a) and (b). The plasmid in lane 2 is identified as pURMM4 and the plasmid in lane 4 is identified as pURMM5.

pURMM4 (Figure 4.12 c, lane 2) and pURMM5 (Figure 4.12 c lane 4) plasmid constructs were expressed in *E.coli* and protein was recovered as inclusion bodies according to the
methods of Ruther & Muller-Hill (1983). pURMM4 and 5 inclusion body preparations were tested for reactivity to FDO 26G (Figure 4.13).

4.6 Western immunoblot analysis of inclusion body preparations from constructs pURMM1-5.

![Western immunoblot analysis of inclusion body preparations from constructs pURMM1-5.](image)

pURMM2, 3 and 5 inclusion body preparations reacted positively with FDO 26G monoclonal antibody, localising the monoclonal antibody FDO 26G epitope to residues 289 and 373 of B3 (3β-HSD type I) protein.

4.7 Peptide mapping experiments.

To further localise the FDO 26G epitope, a series of overlapping synthetic peptides were synthesised that spanned the C-terminal region (Table 4.1). Each peptide included an amino-terminal cysteine for possible conjugation to carrier proteins, and a amino-terminal amide group to minimise charge differences when compared to native protein.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Residues</th>
<th>Amino acid sequence of oligopeptides (single letter code)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5</td>
<td>341-352</td>
<td>NH$_2$-C-K-P-L-Y-S-W-E-E-A-K-Q-K</td>
</tr>
<tr>
<td>C6</td>
<td>346-357</td>
<td>NH$_2$-C-W-E-E-A-K-Q-K-T-V-F-W-V</td>
</tr>
<tr>
<td>C8</td>
<td>356-364</td>
<td>NH$_2$-C-W-V-G-S-L-V-D--R-H</td>
</tr>
</tbody>
</table>
Western immunoblot analysis showed reactivity of FDO 26G antibody with pURMM2 fusion protein at antibody dilution 1:100 and 1:500. Blots containing the same concentration of fusion protein were probed with 1:200 FDO 26G antibody after preabsorption with the test peptides at concentrations of 10 mM. Peptide B6, residues 354-366, absorbed the FDO 26G antibody and resulted in loss of antibody reactivity (Figure 4.14 a).

Peptides were coupled to keyhole limpet haemocyanin (KLH) and tested for reactivity on dot blots. The 354-366 peptide showed significant reactivity, although the shorter internal peptide, 356-364, had very slight reactivity (Figure 4.14 b).

![Western Blot Diagram](image)

**Figure 4.14. Reactivity of FDO 26G with fusion proteins and peptides of 3β-HSD type I.** (a) Western membranes each bearing two lanes of 3β-HSD fusion inclusion body protein from pURMM2. The lower four panels show the pattern of staining when antibody at 1:200 was preabsorbed with each of the peptides (10mM). (b) Reactivity of FDO 26G with three peptides conjugated to keyhole limpet haemocyanin (KLH), at 10, 100 and 1000 ng.

A summary of lacZ gene fusion and oligopeptide mapping data, localising the FDO 26G epitope to residues 354-366 of the B3 (3β-HSD type I) protein is presented in Table 4.2.
4.8. Summary of lacZ gene fusion and oligopeptide mapping data, localising the FDO 26G epitope to residues 354-366 of the B3 (3β-HSD type I) protein.

Table 4.2. Summary of reactivities of FDO 26G with full length 3β-HSD type I, lac Z fusion genes (constructs pURMM1-5), and with synthetic oligopeptides corresponding to the C-terminal region of B3 protein. pURMM1-5 are subclones of B3 cDNA cloned into pUR series of bacterial expression vectors. For each polypeptide/peptide the residue numbers correspond to the amino acid sequence of 3β-HSD type I. The reactivity of each with FDO 26 G is shown as + or -.

<table>
<thead>
<tr>
<th>Construct/peptide residues</th>
<th>FDO 26G epitope</th>
</tr>
</thead>
<tbody>
<tr>
<td>38-HSD type I</td>
<td>+ve</td>
</tr>
<tr>
<td>pURMM1 (1-241)</td>
<td>-ve</td>
</tr>
<tr>
<td>pURMM2 (210-373)</td>
<td>+ve</td>
</tr>
<tr>
<td>pURMM3 (263-373)</td>
<td>+ve</td>
</tr>
<tr>
<td>pURMM4 (241-285)</td>
<td>-ve</td>
</tr>
<tr>
<td>pURMM5 (289-373)</td>
<td>+ve</td>
</tr>
<tr>
<td>peptide 341-352</td>
<td>-</td>
</tr>
<tr>
<td>peptide 346-357</td>
<td>-</td>
</tr>
<tr>
<td>peptide 354-366</td>
<td>+</td>
</tr>
<tr>
<td>peptide 356-364</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Residues 359-362 (S-L-V-D) of the FDO 26G immunoreactive peptide, B6, contain the casein kinase II phosphorylation consensus sequence, (S,T)-X-X-(D,E) with serine 359 the target for phosphorylation. Peptide 354-366 was resynthesised with a phosphoserine at residue 359. The phosphopeptide conjugated to BSA reacted strongly with FDO 26G over a range of dilutions, in a similar way to unphosphorylated peptide. To determine the relative affinities of phosphopeptide and unphosphorylated peptide for FDO 26G, their abilities to preabsorb FDO 26G antibody and extinguish reactivity of 26G was compared to aliquots of non-phosphorylated peptide conjugated to BSA (Figure 4.15).
4.9 Effect of phosphorylation on the binding of FDO 26G to peptide 354-366.

![Table](image)

<table>
<thead>
<tr>
<th>COMPETITOR</th>
<th>PEPTIDE</th>
<th>nM</th>
</tr>
</thead>
<tbody>
<tr>
<td>phosphopeptide</td>
<td>peptide</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>250</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
</tr>
<tr>
<td></td>
<td></td>
<td>750</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10000</td>
</tr>
</tbody>
</table>

Figure 4.15. Peptide 354-366 (unphosphorylated, conjugated to BSA) was immunoblotted onto Hybond membrane. The resulting sixteen pieces of membrane were stained with aliquots of FDO 26G which had been preabsorbed with 0-10,000nM of either phosphopeptide (left hand column) or unphosphorylated (right hand column).

A reduction of antibody binding was evident at concentrations of approximately 250nM free peptide and end points were estimated as 10μM for the unphosphorylated peptide and 1.5 μM for the phosphopeptide (three determinations).
Chapter 5

Characterisation of clone 24-4, a new member of the human 3β-HSD gene family and its identification as a presumptive unprocessed pseudogene (3β-HSDψ1).

5.1 Introduction

Inherited defects of 3β-HSD can result in congenital adrenal hyperplasia (CAH) (Bongiovanni 1962; New and Levine 1984; see section 1.8) and nonsense and frameshift mutations in HSD3B2, the gene encoding 3β-HSD type II, have been described in classical cases of severe CAH affected with salt losing crises, deficiencies in glucocorticoid metabolism and incomplete masculinization (Rheaume et al 1992). A range of less severe phenotypes have also been ascribed to 3β-HSD, including hypospadias, male pseudohermaphroditism, late onset virilisation and clitoral enlargement in women (Pang et al 1985; Mendonca et al 1987; Fisher et al 1987; Frank-Raue et al 1989; New et al 1989).

Four hirsute females from a family exhibiting idiopathic dominant hirsutism were examined. Basal blood levels of Δ5 and Δ4 steroids were within normal limits, but ACTH stimulation led to increases in 17-hydroxyprogrenolone and DHEA. Kindred Phe is shown in figure 5.1.

3β-HSD types I and II genes from affected family members were screened for mutations using PCR amplification and DGGE analysis (Russell 1993). A single base pair substitution, C->A, of base 1426 in the untranslated exon 1 of 3β-HSD type I was identified. No other mutations were detected when the remaining exons were screened, furthermore the haplotype of 3β-HSD types I and II did not segregate with hirsutism in this family (Figure 5.1). However, during mutational analysis by Alan Russell, novel exon 3 and exon 4 fragments were isolated. In particular exon 3 PCR amplification using the intron primer pair (A,B) (Figure 5.2) produced an additional band, (ϕ), in the three members of the Phe family available (Russell 1993).

The ϕ fragment was cloned and sequenced and found to contain an open reading frame, with similar, but not identical sequence to the third exons of the type I and type II genes (Russell 1993). Due to the number of base pair changes it is unlikely that ϕ sequence originated by Taq polymerase errors and is therefore most likely to have come from a distinct gene with similar exon sequence to 3β-HSD types I and II. Furthermore, PCR primers (P,Q) designed to specifically amplify a 124 bp internal fragment of ϕ was present in the genome of unaffected individuals. The appearance of ϕ only in the Phe family with PCR primer pair (A,B) may be explained by point mutations within one of
the intron hybridisation primer sites. It is therefore possible that either mutations or specific rearrangements involving this novel gene may be responsible for the clinical symptoms in the *Phe* family.

To test this hypothesis it was necessary to isolate and sequence the φ gene from one of the patients, and the corresponding gene sequence from an unrelated control individual. PCR primers (P, Q) would allow simple identification of clones containing the φ fragment.

**Family *Phe*: dominant idiopathic hirsutism**

![Family Phe diagram]

Figure 5.1. Family *Phe*. The haplotypes of type I and type II do not segregate with hirsutism in the *Phe* family.
Figure 5.2. PCR amplification and DGGE analysis of HSD3B1 and HSD3B2. Intron/exon organization of HSD3B1 and HSD3B2 and positions of hybridisation sites for PCR and sequencing primers. Primers P, Q hybridise specifically to clone J. S, T were designed to amplify 3β-HSD type II sequences, while the remaining primers hybridise to 3β-HSD type I (see also figure 6.1 a and b).
Results.

5.2 Isolation of the clone 24-4 from a Phe genomic family library and its identification as the φ gene.

In order to isolate the φ gene, a genomic library was constructed from family member IV-2 (Figure 5.1) and hybridised with radioactively labelled B3 (3β-HSD type I cDNA). 14 bacteriophage lambda positive clones were isolated from the Phe library. After a second round of purification, positive plaques were picked as BBL-agarose plugs, and screened by PCR using φ specific primers (P,Q). Clone 24-3 and 24-4 were isolated from the Phe genomic library and supported PCR with (P,Q) (124bp) primers and other PCR primer pairs designed to amplify 3β-HSD type I and II genes (Figure 5.3 a and b).

![Image of gel electrophoresis showing PCR amplification of cloned sequences.](attachment:image.png)

Lane1- 1kb ladder
Lane2- Genomic DNA (L, J)
Lane3- Genomic DNA (T, S)
Lane4- Genomic DNA (A, B)
Lane5- Genomic DNA (P, Q)
Lane6- Genomic DNA (G, R)
Lane7- Clone 24-3 (L, J)
Lane8- Clone 24-3 (T, S)
Lane9- Clone 24-3 (A,B)
Lane10- Clone 24-3 (P, Q)
Lane11- Clone 24-3 (G, R)
Lane12- Clone 24-4 (L, J)
Lane13- Clone 24-4 (T, S)
Lane14- Clone 24-4 (A, B)
Lane15- Clone 24-4 (P, Q)
Lane16- Clone 24-4 (G, R)

![Image of gel electrophoresis showing homologous exon 1/2, exon 3 and exon 4 PCR amplification.](attachment:image.png)

Lane1- 1kb ladder
Lane2- Clone 24-3 (M, K)
Lane3- Clone 24-4 (M, K)
Lane4- Clone 24-3 (L, K)
Lane5- Clone 24-4 (L, K)

Figure 5.3. Homologous exon 1/2, exon 3 and exon 4 PCR amplification of clones 24-3 and 24-4 isolated from the Phe genomic library. a) Clones 24-3 and 24-4 support PCR amplification suggesting that they contain the φ fragment (Lanes 10 and 15). Clone 24-4 is also shown to contain the 3' end of exon 4 (Lane 16). b) Both clones contain homologous exon 1 and 2 sequences (lanes 2-5).
Based upon PCR analysis clone 24-4 contains sequences similar to 3β-HSD exons 1-4, whereas 24-3 does not contain the 3'end of exon 4 (Figure 5.3). Exon 1/2 (PCR primers (M,K)-expected size 437 bp), exon 3 (PCR primers (A,B)-expected size 226 bp) and exon 4 (PCR primers (G,R)- expected size 522bp) PCR products were cloned into T-vector. Transformants were screened using PCR amplification (Figure 5.4 a-c).

Figure 5.4. Clone 24-4 exon homologue PCR products were cloned into T-vector. a) Exon 1/2 transformants were screened using PCR, primers M, K and L, K. b) Exon 3 transformants were screened using PCR primers A, B. c) Exon 4 transformants were screened using PCR, primers G, R.

T-vector clones were identified that contained 24-4 exon homologues; exon 1/2 (Figure 5.4 (a), lanes 4-8), exon 3 (Figure 5.4 (b), lanes 4, 7-9, 11) and exon 4 (Figure 5.4 (c), lanes 4, 5, 7 and 11).

Sequence was determined from T-vector clones containing the 24-4 exon 3 homologue. The sequence, presented in figure 5.13, matched the previously identified φ fragment (Russell 1993), and clone 24-4 was assigned as its corresponding genomic clone.
5.3 Isolation of clone 1-1 from a control genomic library.

A control genomic library was also screened with the 3β-HSD type I cDNA. 43 positive clones reproducibly hybridised through to the secondary stages of purification. Clone 1-1 was identified from this control genomic library and supported amplification with primers (P, Q). Clone 1-1 was also shown to contain exon homologues of a similar size as HSD3B1, HSD3B2 and clone 24-4 (Figure 5.5).

![Image of gel electrophoresis with bands for negative controls, Clone 24-4, and Clone 1-1](image.png)

Lane 1-1 kb ladder  
Lane 2-  
Lane 3- Negative control (M, K)  
Lane 4- Negative control (A, B)  
Lane 5- Negative control (P, Q)  
Lane 6- Negative control (G, R)  
Lane 7- Clone 24-4 (M, K)  
Lane 8- Clone 24-4 (A, B)  
Lane 9- Clone 24-4 (P, Q)  
Lane 10- Clone 1-1 (M, K)  
Lane 11- Clone 1-1 (A, B)  
Lane 12- Clone 1-1 (P, Q)  
Lane 13- Clone 1-1 (G, R)  
Lane 14- Clone 1-1 (G, R)  
Lane 15-  
Lane 16- 1 kb ladder

Figure 5.5. PCR amplification of homologous exons 1/2, exon 3 and exon 4 from clones 24-4 and 1-1. Clone 24-4 was identified from the Phe genomic library, while clone 1-1 was identified from a control human genomic library.

Homologous exon 1/2, exon 3 and exon 4 PCR products from clone 1-1 were cloned into T-vector. Transformants were selected by PCR (Figure 5.6).

![Image of gel electrophoresis with bands for T-vector transformants and 1 kb ladder](image.png)

Lane 1- 20 exon 1/2 T-vector transformants (M, K)  
Lane 2- 48 exon 3 T-vector transformants (A, B)  
Lane 3- 59 exon 4 T-vector transformants (G, R)  
Lane 30- 1 kb ladder  
Lane 60- 1 kb ladder

Figure 5.6. Homologous clone 1-1 exon 1/2, exon 3 and exon 4 PCR fragments were cloned into T-vector. Transformants were screened by PCR; exon 1/2 transformants were screened using PCR, primers (M, K), exon 3 transformants were screened using PCR, primers (A, B), and exon 4 transformants were screened using PCR, primers (G, R).
T-vector clones were identified that contained 1-1 homologous exons; exon 1/2 (Figure 5.6, lanes 1, 3-5, 8-11,13, 15-20), exon 3 (Figure 5.6, lanes 21-23, 25-27, 29, 33, 35, 37, 39-41) and exon 4 (Figure 5.6, lanes 43, 49-51, 54 and 57).

PCR amplification using primer pairs (I,B) and (P,R) (Figure 5.7) yielded products of approximately 4.5kb and 3.5kb respectively. These values were similar with the predicted corresponding amplification products from HSD3B1 and HSD3B2, 4.3kb and 3.9kb (Figure 5.7).

![Image](image.png)

Figure 5.7. PCR amplifications from clones 24-4 and 1-1 showing they contain introns similar in size to HSD3B1 and HSD3B2. Clones 24-4 and 1-1 (I, B) and (A, R) PCR product estimated size approximately 4.5kb and 3.5kb respectively. HSD3B1 and HSD3B2 corresponding PCR products would be predicted as 4.3kb and 3.9kb respectively.

Sequence obtained from 24-4 T-vector clones corresponded exactly with sequence obtained from the 1-1 T-vector clones and had significant similarity with HSD3B1 and HSD3B2 (Figure 5.13). PCR amplification and sequencing was useful for quickly identifying the φ fragment contained within the genomic bacteriophage clones 24-4 and 1-1 and that these clones contained similar sequences to exons 1-4 of 3β-HSD types I and II. However, to complete the sequencing of putative exon material and adjacent potential intron/ exon boundaries, it was necessary to subclone the appropriate fragments from both clones for sequence analysis.

A further objective was to sequence over the (A,B) primer hybridisation sites, to address why the φ fragment was only amplified from the Phe family. It was assumed that subcloning putative exon fragments would allow sequence determination of adjacent intron DNA.
5.4 Southern analysis of bacteriophage clones 24-4 and 1-1.

To identify restriction digestion enzymes useful for subcloning, Southern blots of clones 24-4 and 1-1 lambda DNA were prepared and hybridised with HSD3B1 exon specific probes (Figures 5.8 and 5.9).
Figure 5.8. (a) Southern blot analysis of genomic clones 24-4 and 1-1, isolated from the IV-2 patient and control libraries, and probed with 3B-HSD type 1 (b) exon 2 and (c) exon 3.
Figure 5.9. Southern blot analysis of genomic clones 24-4 and 1-1, isolated from the IV-2 patient and control libraries, and probed with HSD3B1 exon 4.
5.5 Subcloning clone 24-4 putative exons into pUC18.

The 7kb EcoRI fragment of clone 24-4 containing exon 1/2 and exon 3 and the SstI fragment containing exon 4 were subcloned into pUC18 (Figure 5.10 a and b).

![Image](image1)

Lane1- 1kb ladder  
Lane2- Clone 24-4 PCR (A, B)  
Lane3- Negative control (A, B)  
Lanes4- 4-11 exon 3 transformants (A, B)  
Lane 1- lkb ladder  
Lane2- Clone 24-4 PCR (G, R)  
Lane3- Negative control (G, R)  
Lanes4-11 exon 4 transformants (G, R)

Figure 5.10. PCR screening of putative exon 3 and exon 4 subclones from clone 24-4. a) Transformants containing exons 1-3 were screened with PCR primers (A,B) (262 bp). b) Exon 4 transformants were screened using PCR primers (G,R) (520bp).

Subclones that contained 24-4 homologous exons 1-3 (Figure 5.10 a, lanes 4-8 and 10-11) and 24-4 exon 4 (Figure 5.10 b, lanes 10 and 11) were isolated and sequenced (Figures 5.12 and 5.13).

5.6 Subcloning clone 1-1 putative exons into pUC18.

Clone 1-1 putative exon 1/2, exon 3 and exon 4 separate into distinct SstI fragments (Figures 5.8 and 5.9). These fragments were subcloned into pUC18 (Figure 5.11).

![Image](image2)

Lane1- Clone 1-1 bacteriophage DNA restricted with SstI  
Lane 11- 1kb ladder  
Lane2- 8 Clone 1-1 transformants restricted with SstI

Figure 5.11. Transformants containing putative exons 1/2, 3 and 4 were selected by restriction enzyme digestion with SstI and size comparison with 1-1 SstI restricted DNA. SstI fragments containing 1-1 exons had already been identified from Southern analysis, (Figure 5.8).
1-1 subclones that contained similar sequences to HSD3B1 and HSD3B2 were isolated and sequenced (Figure 5.12); exon 1 and exon 2 (Figure 5.11 lane 6), exon 3 (Figure 5.11 lane 7) and exon 4 (Figure 5.11 lane 8). The lineup between 24-4, 1-1, HSD3B1 and HSD3B2 is presented in figure 5.13.

5.7 Sequencing 24-4 and 1-1 subclones.

Figure 5.12. Sequencing extent for clones 24-4 and 1-1. The dotted lines show the sequence obtained from PCR amplified fragments of clones 24-4 and 1-1 after being subcloned into T-vector. The solid lines show the sequence obtained from subclones of the original bacteriophage clones, without the use of PCR. (b) Location of PCR primers in relation to the structure of the 24-4 gene and the restriction enzyme digestion sites Sst I and EcoRI used in the subcloning of clone 24-4 and 1-1.
Exon 2 alignment.

1 50
*                        *
Clone24(4) cttgtcaagt ctataaccac ttacctcctgttttagTC CTGTCATGG
Clone1(1)              ....tg... a. ....t .g. a.- c. ....C .C...G.
Type1                  ........
Type2                  ....ctg... g. ....t ....t ....C .C.T.TG..  
    51 100
* ---Primer I------>*
Clone24(4) TCACCTAGA ATAAGATCTG CTCCCCAGGA TCTTCATATT CCTGGCGAGT
Clone1(1)              ...............
Type1                  ...............
Type2                  ...............

101 150
*                        *
Clone24(4) GTTTCCCTGA ATCTTGGAT- GGCCATGAGT GGCTGGAGCT GCCTTGAGAC
Clone1(1)              ...............
Type1                  ...............
Type2                  ...............

151 200
*                        *
Clone24(4) AGGAGCAGGA GGGTTTCCGG GTCAGAGGAT TGTCAACCTC TTGGTGAAGG
Clone1(1)              ...............
Type1                  ...............
Type2                  ...............

201 250
*                        <=---Primer K---*
Clone24(4) AGAAGGAGCT AAAGGAGATC AGGGCCTTGG ACAAGGCCTC CAGATCAGGA
Clone1(1)              ...............
Type1                  ...............
Type2                  ...............

251 300
*                        *
Clone24(4) TTGAGGGAGA AATTTTCTAg taagtaaact tcagtcgatgg gtgtgtggct
Clone1(1)              ...............
Type1                  ...............
Type2                  ...............

301 350
*                        *
Clone24(4) acattttaaa ccttgcagtg aggtgtggga aggtgagct ttcttgacaa
Clone1(1)              ...............
Type1                  ...............
Type2                  ...............

351 350
*                        *
Clone24(4) gttatg
Clone1(1)              ....t
Type1                  .......
Type2                  ....a.

78
Exon3 alignment

1 50

Clone24(4) ccatccctga acacatatgt aacatcatct ttatcagga aacctcaagc
Clonel(1) ........................................
Type1 ........t..........................c..a..tc..
Type2 t-c..tt..g..........................gc..a..c..

51

Clonel(1) cagattcggga accattcctg atgcgttgcttc gcaagAGCTCC
Type1 ...........at.......................ac a...A...
Type2 ........c.a..at.t........................ac a...

101

Clone24(4) AGAACAAAGAC CAAGATGACA GTGCTAGAA GAGACATTCT GGATCAGTCA
Clonel(1) ........................................
Type1 ...........C..G........................G..C..

151

Clonel(1) TGCCCTGAGA GAGCCTGCCA GGACATCTCG GTAGTCATCC GCACCCCGCTC
Type1 .T.................G.....CA.....A......G
Type2 .T......A..............G.....C.....A......G

201

Clone24(4) TATCATTGAC ATCTTGGTG TCACTCAG AGAGTCTATC ATGAACTTCA
Clonel(1) ........................................
Type1 ........T.G..........................TG..
Type2 ........T.G...T...................C...TG..

251

Clonel(1) ACGTGAAAGg tacaagtgcc tgaagaggag atggagcaag gtaggacaa
Type1 .T.........tg..g...g...........C........g.a.tt
Type2 .T.........g................aa.a....t..tta..t

301 312

Clone24(4) caagagtcatca a
Clonel(1) ............
Type1 a.g.atcaca ..
Type2 g.g.atcaga ..
### Exon 4 alignment

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>acttggagt ggggctgag acacatggat ctgttcaggt gtttggtcgg</td>
<td></td>
</tr>
<tr>
<td>Clone1(1)</td>
<td>.t............g..g. g..a...gttc..c..cacc</td>
<td></td>
</tr>
<tr>
<td>Type1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2</td>
<td>.a..g. g. a...g..t...cagct</td>
<td></td>
</tr>
</tbody>
</table>

| 51 | 100 |  |
| Clone24(4) | ggaagggat atttcctgac attgacagca tgtctttcat gggcagtac |  |
| Clone1(1) | .a...........g..at...........a... |  |
| Type1 |  |  |
| Type2 | .c..t..t..g. |  |

| 101 | 150 |  |
| Clone24(4) | CCCACAGCTG TTGGAGGCT GTGTCCAAGC TAGTGCTACCA GTGCTCCATCC |  |
| Clone1(1) | .tcttt........a...........g........a.t...........a........ |  |
| Type1 |  |  |
| Type2 | .cctttt........g................................ |  |

| 151 | 200 |  |
| Clone24(4) | ATACGAGGTC GCATCGGCTA GCTCTCACAA AGAGATTTTC |  |
| Clone1(1) | .g.tc....a.........................g........a.a.ca.. |  |
| Type1 |  |  |
| Type2 | .g.ta........c.c |  |

| 201 | 250 |  |
| Clone24(4) | CAAATTCAGA ACAACAAGGA GCATCGGCA GACATCTGAC CGTCTCCATA |  |
| Clone1(1) | ..G.C..T.G..G.................C........................T |  |
| Type1 |  |  |
| Type2 | ..G.C..G..G..G................c.ca.. |  |

| 251 | 300 |  |
| Clone24(4) | TCCCCATAGC AAAAAA-CTTG CTGAGAAGGC TGTGCTGGTG GCCTCCTTGGT |  |
| Clone1(1) | C..c...........g...................a..c..c..c.g.. |  |
| Type1 |  |  |
| Type2 | .c..g........c..c |  |

| 301 | 350 |  |
| Clone24(4) | GGACTCTGAA AAATGCCTGG ATGTTGTACA CTTGTCGCTT AAGCTCAATG |  |
| Clone1(1) | .a...........c.gc...c.c..c.............c.ac.c.. |  |
| Type1 |  |  |
| Type2 | .a..a........g..at..c.c.............g..ac..c.ca |  |

<p>| 351 | 400 |  |
| Clone24(4) | CATACCTATG GGGAGGAGG CCCATTTCTT ACTGACAATA TAAAATAGGC |  |
| Clone1(1) | t.............g..c..t..c..t..g..c..g.. |  |
| Type1 |  |  |
| Type2 | t.............g..c..t..c..g..g.. |  |</p>
<table>
<thead>
<tr>
<th>Clone24</th>
<th>Clone1</th>
<th>Type1</th>
<th>Type2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone24(4)</td>
<td>CTTGAAAAAC AATGCAGTCC TGTCAGTCTG TGGAAAACTTC TCCACAAACCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C.C...G... .G... .G... .T.GG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C.C...G... .G... .G...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>ACCCAGCTCT- ------AT GTGGGCTGGG CCCACATTCT GGCCTTTGGA-G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.A TGGTGCA... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.A TGGTGCA.C.. .C... .G...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>GCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAG ACCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>TGAGCAAAGA GTTCGGCCTC TGCCTTGATT CCAGATGGAG GCTTCCTTTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.A...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.A...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>CAGGTCAATT TATACCTATT GACCTCCCTT CAACTGCCAC ACAGTGACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...C... .C... .G... .C... .T...T...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...C... .C... .A... .C...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>TGTCACAATT TATACCTATT GACCTCCCTT CAACTGCCAC ACAGTGACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.A...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.A...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>TGCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>GCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>GCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>GCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>GCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone24(4)</td>
<td>GCCCTGTGGG GCCCCAAGAA GGTCCCCAGGC ATCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type1:</td>
<td>.C...A... .C... .C... .G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type2:</td>
<td>.C...A...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

81
Figure 5.13. Sequence obtained from clones 24-4 and 1-1, aligned with homologous sequences from 3β-HSD type I and II (exons 1-4). The position of PCR primer hybridisation sites are shown (I,K), (A,B), (P,Q) and (G,R). The triplet bases underlined in types I, II and 24-4 are the potential translational start and stop codons. In clone 24-4, exon 3, the underlined sequence is that obtained from sequencing the original PCR band. Dashes indicate gaps introduced to optimise the alignment. Bases corresponding to intron sequence in 3β-HSD types I and II are shown in lower case.
5.8 **Clone 24-4 and clone 1-1 identification as 3β-HSDψ1.**

The codon sequence of the 24-4 and 1-1 genes contained frameshift mutations and premature stop codons mutations within the clone 24-4 and 1-1 putative exon 4 causing premature termination of the predicted polypeptide at amino acid position 173. No other significant open reading frames were found within the sequenced fragments. The first frameshift mutation, deletion of a G residue, was confirmed by sequencing over this region on both strands in clones 24-4 and 1-1 (Figure 5.14).

![DNA sequence from clones 24-4 and 1-1 highlighting the first frameshift mutation](image)

**Figure 5.14.** DNA sequence from clones 24-4 and 1-1 highlighting the first frameshift mutation. a) Clone 24-4 sense strand sequence. This sequence does not contain a HindIII restriction enzyme recognition site because a guanine nucleotide is missing. b) Clone 24-4 reverse strand sequence. c) & d) Clone 1-1 sequence confirms the loss of the HindIII restriction site.
Chapter 6

Characterisation of further novel members of the human 3β-HSD gene family.

6.1 Introduction

Multigene families for 3β-HSD have been described in rodents. To date, five mouse cDNAs (Bain et al 1991; Clarke et al 1993a; Clarke et al 1993b; Abbaszade et al 1995) and four rat cDNAs have been isolated. These isoforms are products of distinct genes, are expressed in a sex- and tissue-specific manner (see section 1.9), and include two functionally distinct groups of proteins. 3β-HSD type I-III in mouse and I, III and IV in rat have dehydrogenase/isomerase activities and utilise NAD+ as cofactor (Bain et al 1991; Zhao HF et al 1991; Keeney et al 1993; Simard et al 1993a). 3β-HSD types IV and V in mouse (Clarke et al 1993b; Abbaszade et al 1995) and type III in rat (DeLaunoit et al 1992) have an exclusive keto-steroid reductase (KSR) activity, which is NADP-dependent and is most likely associated with the inactivation of reduced hormone derivatives such as 5α-dihydrotestosterone. Furthermore the genes encoding the five different isoforms in mouse are closely linked on mouse chromosome 3 syntenic with the human chromosomal region 1p13.

Two forms of 3β-HSD have been described in humans (Luu-The et al 1989; Rheume et al 1991), encoded by genes that are closely linked on chromosome 1p13.1 (Russell et al 1994; Morissette et al 1995; see section 1.9). However, when Southern blots of human genomic DNA are probed with 3β-HSD type I exon 4 DNA, several bands of hybridisation are detected from 0.7kb to 30kb (Russell 1993). Intense bands were observed which corresponded to the sizes predicted by the restriction maps of HSD3B1 and HSD3B2, but other bands of varying size and intensity were also present. It was therefore not surprising that primers designed to amplify fragments of 3β-HSD type I and II genes were found, on occasion, to amplify DNA products of the appropriate length but which resolved as distinct sequence by DGGE analysis (Russell 1993). When exon 3 was amplified using PCR intron primer pair (A,B), two bands were usually detected by DGGE (α and β); in one case a third band (φ; see section 5.1) was detected. When exon 4 was amplified with primer pair (G,H) three additional bands (T, Σ and ρ) were observed on DGGE gels.

Given the multigene family in rodents, Southern data on human genomic DNA and the identification of novel gene segments, it seems likely that the human 3β-HSD gene family mapping to human chromosomal region 1p13 had yet to be fully defined. Some of these gene family members may be unprocessed pseudogenes (Results Chapter 5),
however there is good evidence for 3β-HSD activity in peripheral tissues (Labrie 1991) and for keto-steroid reductase activity in human prostate (Abalain et al 1989).

57 positive lambda clones had previously been isolated from two human genomic libraries that hybridised with 3β-HSD type I cDNA under non-stringent conditions (Results Chapter 5). These clones were screened for novel members of the gene family by PCR amplification followed by DGGE analysis. Various combinations of PCR primers, particularly those that previously amplified novel PCR products from genomic DNA (A,B) and (G,H) (Russell 1993) were utilised (Figure 6.1). It was hoped that genomic clones could be identified that included these fragments as well as other 3β-HSD genes. Also exon 4B PCR amplification would allow rapid analysis of the putative amino acid sequence over the FDO 26G epitope (results chapter 4).

a)

Exon 4B (480bp)

b)

Figure 6.1. Schematic diagram showing the intron/ exon structural organization of (a) HSD3B1 and (b) HSD3B2 (exons I-IV) and positions of hybridisation sites for PCR/ sequencing primers (Not to scale). The position of the the genetic polymorphisms of HSD3B1 in exon 4 and the size of the PCR exon 4B product are indicated, but do not include the GC clamp at one end of the PCR product, required for DGGE analysis. This accounts for an extra 36 base pairs in the exon 4B PCR product.

Initially, the relevant PCR amplification products were cloned into T-vector to allow sequencing. If the clone was confirmed to have novel 3β-HSD sequence, the putative exons were subcloned into pUC18 to allow more extensive sequence analysis.

85
Results

6.2 Analysis of genomic clones by exon 4B (G, H) PCR amplification and DGGE analysis.

3ß-HSD type I gene has two polymorphisms 88bp apart in the coding sequence of exon 4; Bgl II RFLP and the missense (non-synonomous) substitution Asn367Thr. My first piece of analysis is dependent on the exon 4B region (G,H PCR amplification product), therefore the different haplotypes of HSD3B1 must be distinguishable from potential new sequence. Figure 6.2 shows the mobilities of the three most common type I haplotypes; Asn,+, Thr, + and Asn,- in DGGE gels. Also indicated in figure 6.2 are the mobilities of three similar putative exon 4 fragments (T, Σ and ρ) identified during mutation screening by Alan Russell (Russell 1993).

Figure 6.2. DGGE analysis of G, H PCR amplification products from 3ß-HSD type I haplotypes and three homologous exon 4 sequence fragments (ρ, Σ and T) identified from different families (Russell 1993).

Bacteriophage clones were picked as BBL-agarose plugs into 100μl H2O, freeze thawed twice to release bacteriophage DNA suitable for PCR amplification. Primers (G,H) were used for test amplifications from the 57 positive genomic plaques. Primer H was initially designed to hybridise to HSD3B1 as opposed to type II. 35 lambda plaques yielded PCR products of the expected 480bp (Figure 6.1). DGGE was used to distinguish type I haplotypes from novel sequences (Figure 6.3 (a)-(f) and Figure 6.4).
Figure 6.3. Exon 4B DGGE gels of clones isolated from the Phe and control libraries, (a)-(f). PCR amplification of exon 4 using primer pair (G,H). Primer G contains a GC clamp at the 5' end of the primer to improve sensitivity of detection of base pair substitutions within the amplified region. In one gel (d), primer H contains the GC clamp. DGGE classification is indicated in brackets (). (see table 6.1).

a)

Lane 1 - Clone 19-1 (α)
Lane 2 - Clone 19-2 (type I)
Lane 3 - Clone 19-3 (α)
Lane 4 - Clone 19-4 (α)
Lane 5 - Asn + (control)
Lane 6 - Thr + (control)
Lane 7 - Asn - (control)
Lane 8 - Clone 24-2 (type I)
Lane 9 - Clone 23-1 (α)
Lane 10 - Clone 24-1 (α)

b)

Lane 1 - Clone 25-1 (χ)
Lane 2 - Clone 19-1 (α)
Lane 3 - Clone 4-5 (α)
Lane 4 - Clone 2-2 (Asn +)
Lane 5 - Clone 6-2 (χ)
Lane 6 - Asn + (control)
Lane 7 - gap
Lane 8 - gap
Lane 9 - Thr + (control)
Lane 10 - Asn - (control)
Lane 11 - Clone 4-6 (Asn -)
Lane 12 - Clone 4-7 (Asn -)
Lane 13 - Clone 7-1 (Asn -)

C)

Lane 1 - Clone 1-3 (Asn -)
Lane 2 - Clone 1-4 (α)
Lane 3 - Clone 24-2 (α)
Lane 4 - Clone 2-3 (Asn +)
Lane 5 - Clone 2-4 (Asn -)
Lane 6 - 2-6 (χ)
Lane 7 - Asn +(control)
Lane 8 - 3-1 (Asn +)
Lane 9 - 3-2 (χ)
Lane 10 - 3-3 (χ)
Lane 11 - Clone 4-1 (χ)
Lane 12 - Clone 4-2 (Asn +)
Lane 13 - gap
Lane 14 - Clone 4-5 (α)
Although the haplotype controls have not separated well in some cases, there is clear indications that distinct sequences exist in clones identified from the genomic libraries (Figure 6.4).
6.3 Summary DGGE gel of exon 4B analysis.

Figure 6.4. Summary of exon 4B DGGE analysis. Lanes 1-3 contain the exon 4B fragments identified during mutation screening. Lanes 4-6 are type I haplotype controls. The remaining lanes, 7-11, contain exon 4B PCR amplification products from clones isolated from the \textit{Phe} and control genomic libraries.

GH amplification products were resolved by DGGE into at least 4 novel groups; 19-4 (α), 6-5 (γ), 1-2 (p), and 2-1 (T) (Figure 6.3 (a-f) and Figure 6.4). When the genomic clones 1-2 and 2-1 were amplified with PCR primers GH, the PCR product had a similar mobility on DGGE gels as the p and T fragments (Figure 6.4). To confirm these clones contained novel 3\(^{\beta}\)-HSD sequences, further analysis is presented in sections 6.4.1-6.4.4.
6.4 Characterisation of novel clones identified by exon 4B DGGE, from two human genomic libraries.

6.4.1 Clone 19-4. DGGE classification α (Figure 6.4).

Clone 19-4 (G,H) amplification product was cloned into T-vector (Figure 6.5c) and sequenced, confirming that this clone contained novel 3β-HSD sequence. Interestingly, the DNA sequence over the FDO 26G (Results chapter 4), if expressed, would encode a different sequence of amino acids than 3β-HSD type I and II (see Figure 6.13).

Clone 19-4 was shown to have similar structural organization as HSD3B1 and HSD3B2 by PCR (Figure 6.5 a and b). The remaining (M,K) (437bp) and (A,B) (226bp) PCR products were cloned into T-vector (Figure 6.5 c) and sequenced (Figure 6.7 and Figure 6.13) and were also shown to contain novel sequence similar to HSD3B1 and HSD3B2. To allow more extensive sequence analysis of clone 19-4, exon specific Southern blots were used to identify potential restriction enzymes useful for subcloning 19-4 putative exons into pUC 18.

Clone 19-4 putative exons separate into distinct Sst I restriction fragments (Figure 6.6), suitable for subcloning (Figure 6.7).
Figure 6.5. Homologous exon PCR amplifications from clone 19-4 (See Figure 5.2(a) for schematic PCR primer hybridisation sites). (a) 19-4 exon homologue PCR amplifications (1/2, 3 and 4). (b) 19-4 long range PCR using 19-4 phage DNA as template; exons 2-3 across intron B and exons 3-4 across intron C. (c) 19-4 homologous exons 1/2, 3 and 4 cloned into T-vector. Cloning was confirmed by PCR.
Figure 6.6. (a) Restriction analysis of clone 19-4 (α) b) Southern blots of 19-4 DNA hybridised with exon specific probes.
19-4 Sst I restriction products were cloned into pUC18. Clones were shown to contain 19-4 exons by PCR analysis; exon 1/2 (I,K), exon 4 (G,H). The extent of sequencing from T-vector clones and pUC18 subclones is illustrated in figure 6.7.

![Diagram of sequencing extent for clone 19-4. The dotted lines show the sequence obtained from PCR amplified fragments of clone 19-4 after being subcloned into T-vector. The solid lines show the sequence obtained from subclones of the original bacteriophage clones, without the use of PCR. Location of PCR primers in relation to the structure of Clone 19-4 and the Sst I restriction enzyme digestion sites used in the cloning strategy.](image-url)
Clone 6-5 (G,H) PCR amplification product was cloned into T-vector (Figure 6.8).

A number of transformants were identified that contained exon 4 B amplification from clone 6-5. Two transformants (track 5 and 7) were sequenced on both strands. This sequence was identical to the published 38-HSD type II cDNA sequence (Figure 6.15).
6.4.3 Clone 1-2. DGGE exon 4B classification p (Figure 6.4).

Clone 1-2 had a similar mobility on DGGE gels as the previously identified p fragment (Russell 1993). To confirm this result clone 1-2 (G,H) PCR amplification product was cloned into T-vector and sequenced (Figure 6.9).

![Image 1](#)

Figure 6.9. Selection of T-vector clones containing clone 1-2 (G,H) exon 4B amplification product.

A number of transformants were identified that contained clone 1-2 putative exon 4B DNA. Two transformants (lanes 3 and 4) were sequenced on both strands to eliminate the possibility of PCR mutation. The sequence, presented in figure 6.13, matched the p fragment and was assigned as its corresponding genomic clone.

Clone 1-2 was also shown to contain an identical sequence as clone 2-7 within the putative exon 4 region. 2-7 had been identified as a clone containing in the exon 3 DGGE analysis (see section 6.5). A colleague in the laboratory, Alison McVie, under my supervision, was already cloning and sequencing the putative exons of this clone. Clone 2-7 was also identified as the p gene (see Figure 6.10). Therefore no more sequence analysis was performed on clone 1-2. Interestingly the reason clone 2-7 was not identified in the PCR/ DGGE analysis using primer pair (G,H) because clone 2-7 was foreshortened at the 3' end, so it did not contain exon4B sequences. Clone 2-7 sequence, except for the exon 4B region, presented in figure 6.13 was determined by Alison McVie.

6.4.4 Clone 2-1. DGGE exon 4B classification T fragment (Figure 6.4).

Clone 2-1 exon 4B DNA sequence was determined by direct sequencing of the PCR product. It was established that this clone had been misclassified and was shown to contain 3B-HSD type I exon 4B sequence.
6.5 Analysis of genomic clones by exon 3 (A, C) PCR amplification and DGGE analysis.

Of the initial 57 clones, 22 did not amplify with primers GH. 3β-HSD type II exon 3 DGGE PCR primers (A, C) were used to amplify homologous exon 3 fragments from 7 of the 22 clones and were analysed on DGGE gels (Figure 6.10).

The clones amplified using DGGE/PCR primer pair (A,C) were separated into two distinct groups. One group co-migrated with HSD3B2 exon 3 control, while the other group co-migrated with clone 2-7. As has been previously mentioned in section 6.4.3, clone 2-7 had already been assigned as the p gene. Sequence from clone 2-7, except for the exon 4B region, was determined by Alsion McVie (Figure 6.13).
6.6 Identification of further members of the human 3β-HSD gene family by PCR amplification followed by direct sequencing of subsequent PCR products.

Exon 3 and exon 4B PCR and DGGE analysis proved very useful in efficiently grouping 42 of the 57 bacteriophage clones. The remaining 15 clones were classified by PCR amplification using PCR primers hybridising to exons 2, exon 3 and exon 4 of the genes encoding 3β-HSD types I and II, followed by direct sequence analysis. All 15 clones supported PCR amplification with one or more PCR primer pairs. Two further distinct gene sequences were identified (sections 6.6.1 and 6.6.2).

6.6.1 Clone 4-3
Clone 4-3 was amplified using HSD3B2 exon 3 PCR primer pair (D,C) and sequenced (Figure 6.13). This new exon 3 sequence was classified as \( \mu \) (Table 6.1). Further subcloning and sequence analysis of the putative exon2 region was carried out by Nicola Craig (Figure 6.13).

6.6.2 Clone 8-3
Clone 8-3 was initially identified using primers (D,B), sequenced and classified (e) (Table 6.1). Clone 8-3 was shown to contain novel homologous 3β-HSD exons by PCR amplification and direct sequence determination of PCR fragments (Figure 6.11). Limited putative exon sequence was obtained directly from these PCR products in lanes 4 (expected size 280bp), 7 (expected size 235bp) and 9 (expected size 435bp) (Figure 6.13). It was necessary to subclone the putative exon fragments into pUC18 for more extensive sequence analysis. 8-3 bacteriophage DNA was synthesised and analysed on Southern blots (Figure 6.14).
a) Clone 8-3 was shown to contain novel sequence by direct sequencing of the exon 3 PCR product (lane 7). Further test PCR amplifications using a variety of primers amplified fragments of each exon (See Figure 6.1a and b). Exon 3 (lane 7) and exon 4 (lane 9) products were also sequenced. 

b) Sequencing extent for clone 8-3. The dotted lines show the sequence obtained from PCR amplified fragments of clone 8-3 after being subcloned into T-vector. Location of PCR primers in relation to the structure of clone 8-3 and the Sst I restriction enzyme digestion sites that would prove useful in the cloning strategy (Figure 6.12).

Figure 6.11.
Figure 6.14. Restriction and Southern analysis of clone 8-3.  

a) Restriction digestion analysis of clone 8-3. 

b) Southern analysis of clone 8-3 using exon specific probes.
6.7 Summary table of clones and classifications.

Table 6.1. 57 positive clones from two human genomic libraries were separated into distinct gene sequence classifications according to their mobility on DGGE gels and direct sequence analysis of PCR products. Clones 18-1 - 25-2 were isolated from the Phe family library and clones 1-1- 8-3 were isolated from a control human library. Single letter codes refer to PCR/sequencing primers (Figure 6.1 a and b).

<table>
<thead>
<tr>
<th>Clone Number</th>
<th>Classification (Haplotypes of HSD3B1 are indicated where possible)</th>
<th>DGGE analysis</th>
<th>Direct sequencing analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-1</td>
<td>φ</td>
<td></td>
<td>Y,K</td>
</tr>
<tr>
<td>18-2</td>
<td>μ</td>
<td></td>
<td>Y,K</td>
</tr>
<tr>
<td>19-1</td>
<td>α</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-2</td>
<td>Type I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-3</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>19-4</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>19-5</td>
<td>α</td>
<td>G,H</td>
<td>Y,K</td>
</tr>
<tr>
<td>23-1</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>24-1</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>24-2</td>
<td>Type I</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>24-3</td>
<td>φ</td>
<td></td>
<td>A,B and P,Q</td>
</tr>
<tr>
<td>3b-HSDy1</td>
<td>φ</td>
<td></td>
<td>A,B and P,Q</td>
</tr>
<tr>
<td>25-1</td>
<td>α</td>
<td>A,B and P,Q</td>
<td></td>
</tr>
<tr>
<td>25-2</td>
<td>φ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-1</td>
<td>φ</td>
<td></td>
<td>A,B and P,Q</td>
</tr>
<tr>
<td>1-2</td>
<td>α</td>
<td></td>
<td>V,H</td>
</tr>
<tr>
<td>1-3</td>
<td>Type I</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>1-4</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>1-5</td>
<td>Type II (γ)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>1-6</td>
<td>ε</td>
<td></td>
<td>Y,K</td>
</tr>
<tr>
<td>2-1</td>
<td>Type I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-2</td>
<td>Type I (Asn +)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>Type I (Asn -)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>2-4</td>
<td>Type I (Asn +)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>ε</td>
<td></td>
<td>Y,K</td>
</tr>
<tr>
<td>2-6</td>
<td>α</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-7</td>
<td>ρ</td>
<td></td>
<td>A,C</td>
</tr>
<tr>
<td>2-8</td>
<td>ρ</td>
<td></td>
<td>A,C</td>
</tr>
<tr>
<td>2-10</td>
<td>Type II</td>
<td></td>
<td>V,X</td>
</tr>
<tr>
<td>3-1</td>
<td>Type I (Asn +)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>3-2</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>3-3</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>4-1</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>4-2</td>
<td>Type I (Asn +)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-3</td>
<td>μ</td>
<td></td>
<td>D,C</td>
</tr>
<tr>
<td>4-4</td>
<td>α</td>
<td></td>
<td>V,W</td>
</tr>
<tr>
<td>4-5</td>
<td>α</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>4-6</td>
<td>Type I (Asn -)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>4-7</td>
<td>Type I (Asn -)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>5-2</td>
<td>Type II</td>
<td></td>
<td>Y,K</td>
</tr>
<tr>
<td>5-3</td>
<td>Type II (γ)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>5-4</td>
<td>Type I (Asn +)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>5-5</td>
<td>Type II (γ)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td>6-1</td>
<td>Type I (Asn +)</td>
<td>G,H</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>-------</td>
<td>---</td>
</tr>
<tr>
<td>6-2</td>
<td>α</td>
<td></td>
<td>G,H</td>
</tr>
<tr>
<td>6-3</td>
<td>Type I (Asn +)</td>
<td></td>
<td>G,H</td>
</tr>
<tr>
<td>6-4</td>
<td>ρ</td>
<td></td>
<td>D,C</td>
</tr>
<tr>
<td>6-5</td>
<td>Type II (γ)</td>
<td></td>
<td>G,H</td>
</tr>
<tr>
<td>6-6</td>
<td>α</td>
<td></td>
<td>G,H</td>
</tr>
<tr>
<td>6-9</td>
<td>Type II</td>
<td>A,C</td>
<td></td>
</tr>
<tr>
<td>6-10</td>
<td>α</td>
<td></td>
<td>G,H</td>
</tr>
<tr>
<td>7-1</td>
<td>Type I (Asn -)</td>
<td></td>
<td>G,H</td>
</tr>
<tr>
<td>7-2</td>
<td>ρ</td>
<td></td>
<td>D,C</td>
</tr>
<tr>
<td>7-3</td>
<td>ρ</td>
<td>A,C</td>
<td></td>
</tr>
<tr>
<td>8-1</td>
<td>Type II</td>
<td>A,C</td>
<td></td>
</tr>
<tr>
<td>8-2</td>
<td>α</td>
<td></td>
<td>V,W</td>
</tr>
<tr>
<td>8-3</td>
<td>ε</td>
<td></td>
<td>I,J</td>
</tr>
</tbody>
</table>

D,B | V,X |
6.8 Sequence alignment

Figure 6.15. Sequences obtained from clones isolated from two human genomic libraries, aligned with homologous sequences from 3β-HSD type I and type II (exons 1-4). Triplet nucleotides in bold and underlined are translational start and stop codons. 3β-HSD type II missense and nonsense mutations are shown above the aligned sequences in brackets. Corresponding single nucleotides in the clones 19-4, 4-3, 3β-HSD\(\beta\) and 8-3 are highlighted in bold and underlined. The seventeen nucleotides underlined in each sequence are gene specific PCR primers that could potentially be used to map each of the genes to YACs or BACs. Dashes indicate gaps introduced for alignment. Bases corresponding to intron sequence in 3β-HSD types I and II are shown in lower case. Unknown sequence is represented by the character N.

a) Exons 1 and 2
c) Exon 4

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type2</td>
<td>tcttctggg</td>
<td>cagGTACCCA</td>
</tr>
<tr>
<td>Type1</td>
<td>..a..a</td>
<td>..C.....A</td>
</tr>
<tr>
<td>Clone19-4</td>
<td>.t......a.t</td>
<td>.......T..</td>
</tr>
<tr>
<td>Clone2-7</td>
<td>..t..c..g..</td>
<td>.......C.....T..</td>
</tr>
<tr>
<td>3S-HSDv1</td>
<td>....a.........A..AG</td>
<td>.......T..</td>
</tr>
<tr>
<td>Clone8-3</td>
<td>........A......T......A.....T..</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type2</td>
<td>TGTGCCAGTC</td>
</tr>
<tr>
<td>Type1</td>
<td>..A......C.C.</td>
</tr>
<tr>
<td>Clone19-4</td>
<td>.........A.....C.C.</td>
</tr>
<tr>
<td>3S-HSDv1</td>
<td>....T......C.T....C......C.....T.</td>
</tr>
<tr>
<td>Clone8-3</td>
<td>....A.....G.....C.....C.T.....A......T.</td>
</tr>
</tbody>
</table>
251 300

Type2 GTGGCGTTAAG ACCCACATAT ATCTATGGGG AAGGAGGCC ATTCCTTTCT
Type1 .C.C...TG............A...A......
Clone19-4 .C...A.TG.T....A.........T.CA.
Clone2-7 .C...TG.............A......T....
38-HSD\|/1 .C...CT.A.TGC...A...T..A.
Clone8-3 .C...CT...TG......A...A......T.A.A.

301 350

Type2 GCCAGTATAA ATGAGGCCCT GAACAACAAT GGGATCCTGT CAAGTGTTGG
Type1 .T.................C......................
Clone2-7 .A..............A................
38-HSD\|/1 .A..............A...........T...A...........G...............
Clone8-3 .C..............A..............G...............

351 400

Type2 AAAGTCTCTCT ACAGTCAACC CAGTCTATGT TGACCTCACC GCCTGGGCC
Type1 .................C.T..T............T........
Clone19-4 C.................C.G...........A.A........
Clone2-7 C.................C.C...........C........
38-HSD\|/1 C.................C.AC...........G.--.--.--...
Clone8-3 C.................A.C...........A...........T.

401 450

Type2 ACATTCTGGC CTTAGGGGCT CTGGCGGGACC CCAAGAAGGC CCCAAGTGTC
Type1 .................C.A.G.............C.A...........
Clone19-4 C.................C.G...........A.A........
Clone2-7 C.................C.C...........C........
38-HSD\|/1 C.................C..C...........G...............

451 500

Type2 CGAGGTCAAT TCTATTACAT CTCAGATGAC ACGCCTCACC AAAGCTATGA
Type1 .................A.G.............C.T........
Clone19-4 A.G.............C.T...........C........
38-HSD\|/1 T.............A.G...........C........
6.9 Sequence divergence and phylogenetic analysis of the human 3β-HSD gene family.

The relationship between the nucleotide sequences was determined by a weighted parsimony analysis of the alignment between the mammalian 3β-HSD exon sequences and the similar novel sequences of clone 24-4 (ψ1), 8-3 (ψ2), 2-7 (ψ3), 19-4 (ψ4) and 4-3 (ψ5) (Figure 6.16). In figure 6.14, branch lengths represent the number of informative nucleotide substitutions in the data set. The sequence similarity between HSD3B1 and HSD3B2 exon and exon homologues of the new sequences are presented in table 6.2.

Table 6.2. Estimated nucleotide sequence similarity of clone fragments to HSD3B1 and HSD3B2 exons.

<table>
<thead>
<tr>
<th></th>
<th>Type II</th>
<th>Type I</th>
<th>19-4 (38-HSDψ4)</th>
<th>4-3 (38-HSDψ5)</th>
<th>2-7 (38-HSDψ3)</th>
<th>24-4 (38-HSDψ1)</th>
<th>8-3 (38-HSDψ2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type II</td>
<td>-</td>
<td>93.5%</td>
<td>90.5%</td>
<td>89.4%</td>
<td>91.5%</td>
<td>88.2%</td>
<td>85.4%</td>
</tr>
<tr>
<td>Type I</td>
<td>-</td>
<td>91.6%</td>
<td>91.6%</td>
<td>91.5%</td>
<td>88.8%</td>
<td>87.0%</td>
<td></td>
</tr>
<tr>
<td>19-4 (38-HSDψ4)</td>
<td></td>
<td>-</td>
<td>94.5%</td>
<td>91.3%</td>
<td>87.3%</td>
<td>85.4%</td>
<td></td>
</tr>
<tr>
<td>4-3 (38-HSDψ5)</td>
<td></td>
<td></td>
<td>89.4%</td>
<td>90.3%</td>
<td>88.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-7 (38-HSDψ3)</td>
<td></td>
<td></td>
<td>-</td>
<td>88.3%</td>
<td>84.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24-4 (38-HSDψ1)</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>88.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-3 (38-HSDψ2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Figure 6.14. Weighted parsimony analysis of an alignment between expressed mammalian cDNAs and human exon homologue DNA sequences (ψ1-ψ5). Branch lengths represent the number of informative nucleotide substitutions in the data set. Bootstrapping percentages greater than 50% are shown on internal nodes (based on 1000 replications with the input order shuffled randomly on each replication).
The phylogenetic tree in figure 6.14 proved useful in further analysis of the human gene family. Hypothetical ancestors were constructed from the extant genes, which allowed the number of transitions and transversion to be characterised throughout the evolution of the macaque and human gene family (Table 6.3 and Table 6.4).

Table 6.3. Numbers of transitions and transversions within the primate lineage from known sequences; 1st, 2nd and 3rd nucleotides of the codon are indicated. Hypothetical ancestors are highlighted in BOLD. Ancestor 6 is included in both tables, because from this hypothetical gene sequence arose both the active and three of the unprocessed pseudogenes.

<table>
<thead>
<tr>
<th></th>
<th>Transitions</th>
<th></th>
<th>Transversions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>Total</td>
</tr>
<tr>
<td>Macaque</td>
<td>3</td>
<td>1</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>Anc-II</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Type II</td>
<td>3</td>
<td>5</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>Anc-Primates</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Type I</td>
<td>8</td>
<td>3</td>
<td>16</td>
<td>27</td>
</tr>
<tr>
<td>Anc-6</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Transitions</th>
<th></th>
<th>Transversions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>Total</td>
</tr>
<tr>
<td>Anc-6</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>19-4</td>
<td>12</td>
<td>8</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>Anc 19-4/4-3</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>4-3</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Anc 19-4/4-3/2-7</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>2-7</td>
<td>9</td>
<td>10</td>
<td>13</td>
<td>32</td>
</tr>
<tr>
<td>24-4</td>
<td>18</td>
<td>9</td>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>Anc 24-4/8-3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>8-3</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>175</td>
</tr>
</tbody>
</table>

111
Table 6.4. Types of nucleotide substitutions, within the exons of functional primate genes and similar sequences in the pseudogenes. a) The active genes include human 3β-HSD type I and II and macaque. b) The unprocessed pseudogenes are the novel human sequences identified in results chapters 5 and 6. c) Final nucleotide ratios in the exons of the functional genes and similar sequences in the pseudogenes.

### a) Functional genes.

<table>
<thead>
<tr>
<th>From</th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-</td>
<td>5</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>-</td>
<td>5</td>
<td>31</td>
</tr>
<tr>
<td>G</td>
<td>26</td>
<td>7</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>22</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

Transitions - 94 (75.2%)
Transversions - 31 (24.8%)

### b) Pseudogenes.

<table>
<thead>
<tr>
<th>From</th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-</td>
<td>12</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>19</td>
<td>-</td>
<td>11</td>
<td>65</td>
</tr>
<tr>
<td>G</td>
<td>56</td>
<td>19</td>
<td>-</td>
<td>16</td>
</tr>
<tr>
<td>T</td>
<td>11</td>
<td>28</td>
<td>6</td>
<td>-</td>
</tr>
</tbody>
</table>

Transitions - 175 (63.2%)
Transversions - 102 (36.8%)

### c) A+T ratios C+G ratios CpG dinucleotide frequencies

<table>
<thead>
<tr>
<th></th>
<th>A+T ratios</th>
<th>C+G ratios</th>
<th>CpG dinucleotide frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macaque</td>
<td>47.9%</td>
<td>52.1%</td>
<td>19</td>
</tr>
<tr>
<td>Type II</td>
<td>49.0%</td>
<td>51.0%</td>
<td>22</td>
</tr>
<tr>
<td>Type I</td>
<td>48.9%</td>
<td>51.1%</td>
<td>28</td>
</tr>
<tr>
<td>19-4 (ψ4)</td>
<td>48.8%</td>
<td>51.2%</td>
<td>19</td>
</tr>
<tr>
<td>4-3 (ψ5)</td>
<td>N.D</td>
<td>N.D</td>
<td>N.D</td>
</tr>
<tr>
<td>2-7 (ψ3)</td>
<td>49.7%</td>
<td>50.3%</td>
<td>16</td>
</tr>
<tr>
<td>24-4 (ψ1)</td>
<td>51.0%</td>
<td>49.2%</td>
<td>12</td>
</tr>
<tr>
<td>8-3 (ψ2)</td>
<td>53.8%</td>
<td>46.2%</td>
<td>8</td>
</tr>
</tbody>
</table>
Chapter 7

General discussion

7.1 Mapping the monoclonal antibody FDO 26G epitope.

3β-HSD enzyme activity is present in the microsomes of chorion laeve at levels similar to those in villous syncytiotrophoblast (Gibb et al 1978, Gibb et al 1985), and chorion laeve can synthesise progesterone from labelled pregnenolone (Mitchell et al 1982). This is consistent with monoclonal antibody FDO 161G reactivity in these placental tissue (Mueller 1987). The expression of 3β-HSD in intrauterine trophoblast indicates a capacity to synthesise Δ4- steroids from circulating precursors. Androstenedione can then be converted to oestrogens by placental aromatase (Kilgore et al 1992). Thus, intrauterine trophoblast is likely to contribute to the hormonal maintenance of the uterine decidua in early pregnancy. FDO 161G reacts in a tissue specific manner to 3β-HSD protein in syncytiotrophoblast, nonvillous trophoblast, sebaceous glands in skin, cells of the adrenal cortex and the gonads. This is a similar pattern of reactivity to that described by others using polyclonal antisera (Pelletier et al 1992; Riley et al 1992). Since expression of 3β-HSD type I and II are almost mutually exclusive in placenta and adrenals, it seems likely that FDO 161G binds a common epitope in the two forms. This epitope is probably conformational, being undetectable in western blots. Monclonal antibody FDO 26G was raised against 3β-HSD type I which had been purified from villous syncytiotrophoblast, the sequence of which had been confirmed by sequencing an N-terminal and an internal peptide (Nickson et al 1991). This antibody showed a more restrictive staining pattern; reacting strongly with villous syncytiotrophoblast and adrenal cortex, but their was greatly reduced reactivity with trophoblast from the chorion laeve and no reactivity with intrauterine trophoblast and cytotrophoblast cell islands. These staining patterns were unaffected by varying dilutions of the monoclonal antibodies over a wide range, and we infer that the lack of FDO 26G reactivity with non-villous trophoblast is due to the expression of a different isoform of the enzyme. When sheep polyclonal anti-3β-HSD was used to stain western blots, protein bands of the same mobility were detected in villous trophoblast and chorion laeve (Hawes et al 1994) and it may have been that same fraction that was detected in western blots.

To identify the cDNA encoding FDO 161G reactive protein, a previously identified, partial cDNA fragment, I/6, was used to screen a λ gt11 cDNA expression library, in which 48 positive clones were identified. Ten positive clones were selected randomly for purification and further analysis. cDNA insert size was determined by PCR amplification using primers hybridising to the λgt11 bacteriophage arms. B3 (P10) was shown to contain an insert of approximately 1.7 kb (Figure 3.2), was sequenced and identified as 3β-HSD type I (Nickson et al 1991). It was been demonstrated, by immunofluorescent
(FITC) staining, that clone B3 contains an open reading frame encoding a polypeptide that transiently expressed in COS-7 cells reacts with the monoclonal antibody FDO 161G (Figure 3.8). B3 was also stably transfected into Chinese hamster ovary (CHO) cells and reactivity was confirmed to the monoclonal antibody FDO 26G by western immunoblot analysis (Figure 3.9). Given the immunostaining patterns described in section 1.13 from our colleagues in Australia, it seemed likely that both monoclonal antibodies react with the same protein in vivo and that the discrepancies in staining patterns may either be explained by post-translational modification or by the presence of an unidentified member of the 3β-HSD gene family, that may cross react with FDO 161G but not FDO 26G. Analysis of the remaining clones isolated from the placental library using denaturing gradient gel electrophoresis (DGGE) confirmed that all clones identified in the screen, which reproducibly hybridised with 1/6 DNA, were haplotypes of 3β-HSD type I (Figure 3.12, summarised in Table 3.1). This suggests that 3β-HSD type I is the major transcript expressed in the placenta. However, the amplification of the cDNA library may favour one form of 3β-HSD over another, therefore we cannot exclude the possibility of expression of other 3β-HSD isoforms in the placenta based on this data alone.

The next line of investigation was to locate and identify the FDO 26G epitope. Lac Z fusion genes and oligopeptide mapping located the FDO 26G epitope to the amino acid sequence VEWVGLVDRHKE, residues 354-366 of 3β-HSD type I (Figure 4.14). The same sequence exists in 3β-HSD type II (residues 353-365) consistent with FDO 26G reactivity in the adrenal cortex. The epitope contains a consensus for casein kinase II (CK-II) phosphorylation site (S,T X X D,E). The glutamic acid residues at -4 and +4 from the consensus are associated with CK-II activation, as is the choice of S-359 and D-362 at the consensus residue (Pinna 1990). Nevertheless, the peptide competition data (Figure 4.15) shows that phosphorylation of S-359 has only a marginal effect on the binding of peptide to antibody, with evidence of the phosphopeptide having an higher affinity. This data would suggest that differential phosphorylation of S-359 could not account for the extinction of FDO 26G reactivity on extravillous trophoblast. There were no other identifiable post translational modification consensus sequences within the FDO 26G epitope (residues 354-366), therefore no other putative modifications were tested. More recently, the staining data from adult Leydig cells, which reacts with FDO 161G but not FDO 26G, supports the post-translational modification hypothesis. At present the staining pattern of these antibodies in fetal Leydig cells is unknown, but it would be interesting to investigate the role of post-translational modification of 3β-HSD during human development. Figure 7.1 shows an alignment of amino acid sequences homologous to the human 3β-HSD type I peptide, over the FDO 26G epitope (residues 354-366). The human type I is
represented in its two allelic forms. All mammalian 3β-HSD enzymes have diverged over the FDO 26G epitope except for the human type I, II and macaque 3β-HSD. The four sequences for rat 3β-HSD have diverged from the type I human peptide sequence and no reaction was observed of FDO 26G with frozen sections of rat steroidogenic tissues (U. M. Mueller personnel communication). Despite this divergence it is notable that the potential casein kinase II phosphorylation site (CK-II) is widespread amongst mammalian 3β-HSD. A further CK-II consensus site is located 14 residues upstream in all these species (human type I, residues 344-347: SWEE). A protein kinase C (PKC) consensus site is located seven residues from the carboxy terminus of 3β-HSD from bovine, human and macaque but not rodents. The basic lysine residues at -2 and +2 from the consensus site are associated with the enhancement of functional PKC sites (Woodget et al., 1986). A genetic polymorphism exists at residue 367 (Russell et al., 1991, 1994). Since clinically normal individuals can readily be identified who are homozygous for 367<sup>N/N</sup> allele, the putative PKC site in the type I protein cannot fulfil a vital function. No such polymorphism exists in the human type II isoform, which is expressed in the adrenal cortex and gonads, and may be subject to different regulatory circuits.

**Figure 7.1.** The C-terminal 20 amino acid residues from various forms of mammalian 3β-HSD, showing the position of the FDO 26G epitope in the human enzyme and the consensus site of phosphorylation by the protein kinases, casein kinase II (CK-II) and protein kinase C (PK-C). The arrow marks the position of amino acid residue 367 in the human 3β-HSD type I enzyme. * represents the stop codon.
Protein kinase A (PK-A) is the major signal regulating transcription of 3β-HSD in response to the action of trophic hormones at the cell surface, via cyclic AMP (cAMP). There is also evidence for the interactive effects of PKA and PKC stimulation on the levels of 3β-HSD mRNA in choriocarcinoma cells (Tremblay & Beaudoin 1993). In cultured human adrenocortical cells, such interactions lead to increases in 3β-HSD activity, elevation of non-17α-hydroxylated steroids and reduction of adrenal androgen synthesis (McAllister & Horsby 1988). CK-II phosphorylation is observed in many classes of protein, though the functional consequences of this remain to be established (Pinna 1990). CK-II acts synergistically with glycogen synthase kinase-3 stimulating the phosphorylation of phosphoprotein phosphatase (DePaoli-Roach 1984), and with PK-A on dopamine- and cAMP-regulated phosphoprotein (Girault et al. 1989). Consensus sites for CK-II, PK-A and PK-C have been found clustered at the N-terminus of the cAMP response element-binding protein from PC12 cells (Gonzalez et al. 1989). It is therefore possible that the CK-II and PK-C consensus sites at the C terminus of the 3β-HSD polypeptide are involved in either the post-transcriptional regulation of 3β-HSD activity or substrate specificity, potentially modifying or amplifying the effect of transcriptional regulation.

7.2 Investigation into the extent of the human 3β-HSD gene family.

The association between P450c21 deficiency and hirsutism led to interest in the possible role of mutations in 3β-HSD in the pathogenesis of hirsutism. However, deficiency of 3β-HSD type I has yet to be reported and proven deficiency of 3β-HSD II is rare. Only a small number of females have been reported with type II deficiency, some show clinical phenotypes, including amenorrhea, virilization, and premature pubarche (Mendonca et al. 1994, Sanchez et al. 1994b, Paula et al. 1995). In others the mutation has no clinical effect despite markedly abnormal blood steroid levels (Mendonca et al. 1994) and heterozygotes for partial 3β-HSD deficiency are not affected with hirsutism (Russell et al. 1994). Mathieson et al. (1992) failed to find evidence of 3β-HSD deficiency in a large series of patients affected with primary hirsutism, but others have claimed that a non-classical form of 3β-HSD deficiency existed in hirsutism (Pang et al. 1985, Zerah et al. 1991). This deduction was based on serum ratios of Δ5/Δ4 steroids, and the effect of ACTH stimulation, the results of which exceeded the normal range, but were less extreme than was found in cases now proven to be affected with partial 3β-HSD type II deficiency. Recently, these biochemical signs were re-investigated in a group of patients previously classed as having non-classical 3β-HSD mutations (Zerah et al. 1994). The biochemical signs were found to be inconsistent over time and no mutation were detected in the genes for 3β-HSD type I and II.
In the *Phe* family, hirsutism was found in seven out of eleven family members over four generations, with a pattern of sex-limited dominant inheritance (Figure 5.1). Kindred *Phe* was not suggestive of classical or non-classical 3ß-HSD deficiency and was included in an ongoing screen for mutations in 3ß-HSD because of the intriguing dominant mode of inheritance and the evidence of increased responsiveness of 17-hydroxyprogrenolone and DHEA to synacthen. There was no basal elevation of Δ5 steroids, nor a general secondary rise in Δ4 steroids due to peripheral steroid metabolism. The haplotypes of 3ß-HSD types I (Asn-; Asn+) and II (intron C microsatellite) did not cosegregate with hirsutism in case IV-5, indicating that, unless IV-5 was a sporadic case (unlikely since III-5, the mother of IV-5 was also affected) of unrelated aetiology, hirsutism in this family was not closely linked to the genes for 3ß-HSD types I and II (HSD3B1 and HSD3B2 respectively), which map to 1p13.

During the routine screening for mutations in the genes for 3ß-HSD types I and II, a novel homologous exon 3 fragment (ϕ) containing an open reading frame was identified from a patient IV-2 (Figure 5.1) suggesting that, in some families, the expression of other human 3ß-HSD genes may be responsible for 3ß-HSD deficiency. Corresponding genes were isolated from patient IV-2 and control human genomic libraries, clones 24-4 and 1-1 respectively (sections 5.2 and 5.3). The order of the putative exons for clones 24-4 and 1-1 were confirmed by PCR (Figure 5.7) and Southern blot evidence (Figures 5.8 and 5.9) also confirms that clones 24-4 and 1-1 are very similar at the structural level. Therefore it is unlikely there is any gross rearrangement within clone 24-4, that could account for the phenotype in the *Phe* family. The sequence of the putative exons and adjacent intron fragments of the genomic clones 24-4 and 1-1 were determined, and found to be identical, including the sequences that flanked exon 3 from which band ϕ was amplified (PCR primer hybridisation sites (A,B)(Figure 5.13). It is difficult to explain why the ϕ fragment was initially only amplified in the *Phe* family using PCR primer pair (A,B). The presence of nonsense and frameshift mutations in clones 24-4 argues that it is not expressed as a functional 3ß-HSD enzyme, and is probably an intron containing unprocessed pseudogene which has been identified as 3ß-HSDψ1 (McBride et al 1995a).

Using rat- human somatic cell hybrids, (P,Q) PCR primer pair were used to map the novel exon 3 fragment to chromosome 1. Using the 3ß-HSDψ1 genomic clone, the map position was further refined by fluorescent in situ hybridisation (FISH) to the human chromosomal region 1p13 (McBride et al 1995). The location of 3ß-HSD ψ1 at a similar chromosome location as genes encoding the enzymes 3ß-HSD type I and II enzymes, also argues against the involvement of 3ß-HSDψ1 in the *Phe* family phenotype, though the linkage between the genes for types I and II and 3ß-HSDψ1, at this time, had yet to be established. Interestingly, a cluster of functional 3ß-HSD genes, including 3ß-HSD
keto-steroid reductase genes, has been reported on mouse chromosome 3 (Bain et al 1993), syntenic with the human chromosomal region 1p13. This suggests that other 3β-HSD gene sequences in addition to 3β-HSD ψ1 may exist at this location in the human. The size of the cluster has yet to be fully defined, though it may not include the gene for 3β-hydroxy-Δ5-C27-steroid dehydrogenase (Russell et al. 1995b).

The findings from Southern blots hybridised with 3β-HSD type I exon 4 provides evidence of the existence of further members of the human 3β-HSD gene family, but it was not possible to deduce how many different members of the family existed. In an attempt to define the extent of the 3β-HSD family in humans, 57 clones were isolated from two genomic DNA libraries by homology probing with 3β-HSD type I cDNA. Analysis of these clones by non-stringent PCR amplification, DGGE, and PCR sequencing led to the identification of five novel gene sequences similar to HSD3B1 and HSD3B2 (Figure 6.13). The groups identified were; HSD3B1 (15 clones, roughly equal proportions of the HSD3B1 haplotypes Asn- and Asn+), HSD3B2 (8 clones), 19-4 clones (18 clones), 4-3 (2 clones), 2-7 (6 clones), 3β-HSDψ1 (5 clones) and 8-3 (3 clones)(Table 6.1). The difference in representation of clones may not be significant. The numbers will reflect efficiencies of cloning different sequences, and a proportion of sibling clones are likely to have been isolated since the libraries had been amplified before this study. All novel genomic clones were mapped initially using FISH, to the human chromosomal region 1p13. With the identification of gene specific PCR primer pairs (Figure 6.13), it was possible to undertake a more detailed mapping strategy, to localise and order the known 3β-HSD sequences to yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs) (McBride et al 1997).

It is difficult to predict whether there are as yet unidentified 3β-HSD functional genes present in the human genome. Two inconsistencies were noted during the DGGE analysis that may suggest the presence of further genes. Clones initially designated χ; Figure 6.3 (b) lane 1 (clone 25-1), lane 5 (clone 6-2) and Figure 6.3 (c) lane 6 (clone 2-6), lane 9 (clone 3-2), lane 10 (clone 3-3) and lane 11 (clone 4-1), were subsequently shown to have the same mobility as groups of clones designated α; Figure 6.3 (f) lane 8 (clone 3-3), lane 10 (clone 2-6), lane 11 (clone 6-2), lane 12 (clone 25-1) and Figure 6.4 lane 10 (clone 3-3). Mutation introduced by the PCR amplification reaction cannot explain this inconsistency because independent PCR products are unlikely to contain the same PCR mutations (see Figure 6.3 b and c). It is possible that initially the DGGE gels detected a nucleotide substitution within these fragments that was not seperated in subsequent DGGE gels. Further DGGE analysis could have been undertaken using a 3’ GC clamp primer or sequence of the exon 4B PCR product from these inconsistent clones could have been determined. Due to other priorities these experiments were not undertaken. In
the HSD3B2 exon 3 DGGE analysis (Figure 6.10) extra putative exon 3 fragments were observed, however, as a full exon 3 DGGE analysis was not carried out on the genomic clones, it is uncertain whether these extra bands represent unidentified exon 3 sequences or exon homologues in the other newly identified members of the 3β-HSD gene family.

An alignment of 3β-HSD types I and II with exon homologues identified from each of the novel genes is presented in figure 6.13. It is of interest that a number of these sequences, if expressed would encode different amino acid sequences over the FDO 26G epitope. However, although 3β-HSDψ1, and the four new sequences; 8-3, 2-7, 19-4 and 4-3 (Figure 6.13) all have ATG codons in equivalent position to the start codon of the type I gene, all have stop codons that occur well before the stop codons of the type I and II genes (clone 24-4 has a stop at residue 173; clone 8-3 a stop at residue 143; 2-7 has a stop at residue 42; and clones 19-4 and 4-3 have a stop at residue 16) as well as other missense and frameshift mutations. It is clear that the deletion of any major portion of 3β-HSD I or II would lead to an inactive enzyme. Exon 2 of 3-HSD I and II contains a nucleotide binding (Wierenga et al.1986) consensus sequence, and substrate protection studies show the substrate binding site of 3β-HSD type I is located in exon 4 (between residues 175 and 274) (Thomas et al.1993). Further, type II mutations have been detected between residues 15-318, and all either abolish or greatly reduce enzyme activity (see Table 1.1). Since the integrity of all three exons is required, these stop codons would disrupt enzyme activity. RNA products from 19-4 and 4-3 were detected by reverse transcription with oligo (dT) and therefore had been processed with poly (A) tails (S. Burridge personnel communication). However, these products showed forms of alternate splicing that created frame shifts from exon 3 to 4 and/or inserted intron sequence. This suggests that these sequences do not encode functional 3β-HSD enzymes and therefore should classified as unprocessed pseudogenes, which we name 8-3 (ψ2), 2-7 (ψ3), 19-4 (ψ4) and 4-3 (ψ5). 3β-HSDψ1 was identified in results chapter 5. Presumably these sequences have been generated by gene duplication events.

Gene duplication is a critical process in the evolution of higher organisms. It recruits pre-existing genetic material as a substrate for the formation of novel functional units. Extra gene copies created through duplications may ultimately diverge to perform related specialised developmental and biochemical function. This process is exemplified by the globin family which has evolved a highly co-ordinated process of tissue- and stage-specific expression of developmentally specialised genes from a single ancestral gene (Goodman 1987). The apparent necessity for the differential gene expression during development presumably has provided selective pressure to maintain expression of multiple copies. Genes can duplicate by a variety of mechanisms (Maeda & Smithies 1986). How much of the pattern of duplication that can be unravelled in any given gene
family generally depends on the age and complexity of the family. Initial duplication events of a single copy gene are rare, but may be catalysed by repetitive elements such as Alu or L1 (Cross and Renawitz 1990). Depending on the number of genes and the location of the cross-over events, duplication by this mechanism can drastically affect the gene copy number of a family. After the gene has duplicated, the two copies start to diverge by accumulating mutations. If there is no selective advantage for multiple copies, one of the genes could either develop different regulation from the original copy, become a non-functional pseudogene, or fortuitously gain a new function. Presumably, because of the similarity between the genes, recombination may be frequent.

In the alignment of human 3β-HSD in figure 6.13, the new pseudogene sequences diverge at a number of positions including missense substitutions that correspond to 3β-HSD type II mutations. This raises the question of whether certain DNA mutations associated with 3β-HSD type II deficiency may have arisen from gene conversion events similar to those suggested in the molecular genetics of P450c21 deficiency, encoded by CYP21 (see section 1.8). Although mechanism of apparent gene conversion in higher eukaryotes is not well understood, there are strong reasons for suspecting it occurs. In P450c21 deficiency apparent gene conversion events generate CYP21 mutations when short sequences from the highly homologous (98% over exon homologues) CYP21 pseudogene are transferred to the active gene. Collier 1993 suggests that a gene conversion involving the exchange of up to 390 nucleotides arose by a microconversion event between CYP21 and CYP21P. Intergenic recombination has also been shown to be responsible for glucocorticoid-remediable aldosteronism and human hypertension (Lifton et al. 1992). The tandem arrangement of the closely related CYP11B1 and CYP11B2 genes results in an unequal crossing over, leading to a chimeric gene expressing aldosterone synthesising properties under the regulatory control of ACTH (Figure 1.6).

Although more recent investigations into detection of mutations within the CYP11B1 gene suggest that replication based mutagenesis is more likely to account for CYP11B1 deficiency than either gene conversion or recombination events (Skinner et al. 1996).

Turning to the present data (Figure 6.13), above the type II exon sequence are shown positions of 22 mutations in the type II gene that have been observed in cases of congenital adrenal hyperplasia, pseudohemaphroditism and premature pubarche published to date. In 17 (77%) of these mutations the type II mutations is not present in 3β-HSD ψ1-ψ5, and cannot be the source of these mutations. The remaining 5 (23%) mutations are present in 3β-HSD pseudogene sequence, highlighted in bold and underlined (Figure 6.13). However, type II mutations in 3β-HSD deficiency mainly have single nucleotide substitutions. If gene conversion did occur we would expect to see a number of nucleotide substitutions reflecting the nucleotide difference between 3β-HSD type II and the donar pseudogene sequence. Such multiple mutations would be expected
to have severe effects on type II enzyme function as in cases of classical salt wasting congenital adrenal hyperplasia. To date no such mutations have been described. It is probable that the divergence between the members of the 3\(\beta\)-HSD gene cluster is sufficient to reduce the occurrence of conversion and recombination, as well separated point mutations efficiently inhibit homologous recombination (Leach 1996).

The relationship between the nucleotide sequences was determined by a weighted parsimony analysis (Figure 6.14) of the alignment between the mammalian 3\(\beta\)-HSD exon sequence and the homologous sequences of 3\(\beta\)-HSD\(\psi\)1-5. In figure 6.14, branch lengths represent the number of informative nucleotide substitutions in the data set. The human sequences for type I, II, and pseudogenes are more closely related to each other than to rat or mouse. As previously noted (Abbaszade et al 1995), this indicates that the human gene family expanded since the divergence from the common ancestor with rodents. Human pseudogenes 3\(\beta\)-HSD\(\psi\)1-5 have ATG in the homologous position to the initiator ATG codon of the type I gene, which is placed one codon earlier than the type II gene. This suggests that they are more closely related to the type I gene, a view supported by the analysis in figure 6.14. Pseudogenes \(\psi\)4 and \(\psi\)5 are both expressed as RNA in several tissues (S. Burridge personnel communication) and show the closest homology to the type I and II genes (Table 6.2). This suggests that they either arose more recently than the other pseudogenes, or have been more recently released from the selective pressure associated with the expression of functional enzymes. There is insufficient evidence to determine whether \(\psi\)4 and \(\psi\)5 represent extinct enzymatic functions distinct from type I and II, although it is possible to date the duplication event that gave rise to these genes, because they contain the same inactivation mutation (Figure 6.13). Assuming that noncoding regions of human DNA accumulate at the rate of approximately 0.3%/ million years (Maeda et al 1996) and given the percentage divergence of \(\psi\)4 and \(\psi\)5 (Table 6.2), this dates the duplication event to approximately 18 million years ago.

Further analysis suggests there was rapid amplification of the human gene family, because of the low number of transitions and transversions in the hypothetical ancestors calculated from the extant gene sequences (Table 6.3). The non-random pattern of nucleotide substitutions within the 3\(\beta\)-HSD unprocessed pseudogene family are characteristic of pseudogene sequences (Li et al 1984). Nucleotide substitutions occur more frequently in the first and second positions of the 'codon' within pseudogenes (Table 6.3). Transitions are more likely to occur, because transversions require two rare tautomeric and rotameric states of the bases to form a purine-purine mispair that avoids detection by the repair systems, whereas transitions require only one (Topal & Fresco 1976). The rates of both transitions and transversions collectively lead to a gradual decrease in the (G+C) content (Table 6.4). The reason for this gradual change is not because of inherent liability in the CpG dinucleotide, but because cytosine residues 5' to neighbouring guanines are prone to
be methylated. 5-methyl-cytosine is more susceptible to spontaneous deamination, which leads to the formation of thymine, a normal residue in DNA. Generally, the CG dinucleotide occurs in low frequencies in metazoan DNAs (Ehrlich et al 1990), whereas we do not see this in E.Coli DNA, which is not subject to the same pattern of methylation.

In rodents, ketosteroid reductase and dehydrogenase/isomerase activities are due to distinct members of the 3β-HSD gene family, which had probably diverged in a common ancestor to rat and mouse (Abbaszade et al 1995). In human, both activities can be detected when either the type I or the type II enzyme are expressed in transfected tissue culture cells (Rheaume et al 1991; Lorence et al 1990). This apparent dual function of the human enzyme may account for the presence of ketosteroid reductase activity in human prostate, since that tissue contains detectable levels of 3β-HSD type II RNA (S. Burridge personnel communication). The dual function of the two human 3β-HSD enzymes may explain why the human genome apparently has a small number of functional 3β-HSD genes, compared with rodents. However, the lack of a specific ketosteroid reductase gene in man has not been specifically proven in this study; such a gene may have been underrepresented in the library, or be sufficiently diverged in sequence to be undetected when the library was screened.

Surprisingly the present work suggests independent expansion of the human 3β-HSD gene family since the divergence from the common ancestor with rodents and bovine. The expectation was that expansion of the 3β-HSD gene family occurred in a common ancestor of primates and rodents. However, in the macaque (old world monkey), it seems likely that the gene family is also expanded (Figure 6.14) raising the question whether more than two functional genes are expressed in this lineage, as there is clear evidence for expression of 3β-HSD enzymes in peripheral tissues such as liver and kidney. Further investigations into the 3β-HSD gene families within primate lineages may yield important clues into the timing of expansion of the 3β-HSD gene family.


Baker ME, Blasco R (1992). Expansion of the mammalian 3-beta-hydroxysteroid dehydrogenase plant dihydroflavonol reductase superfamily to include a bacterial cholesterol dehydrogenase, a bacterial UDP-galactose-4-epimerase, and open reading
frames in vaccinia virus and fish lymphocystis disease virus. Febs Letters 301: 89-93.


Cidlowski JA, Cidlowski NB (1982). Glucocorticoid receptors and the cell-cycle -
evidence that the accumulation of glucocorticoid receptors during the s-phase of the
cell-cycle is dependent on ribonucleic-acid and protein-synthesis. Endocrinology
110: 1653-1662.

Clarke TR Bain PA, Greco TL, Payne AH (1993b). A novel mouse kidney 3-beta-
hydroxysteroid dehydrogenase complementary-dna encodes a 3-ketosteroid
reductase instead of a 3- beta-hydroxysteroid dehydrogenase/delta(5)-delta(4)-
isomerase. Molecular Endocrinology 7: 1569-1578.

Clarke TR, Bain PA, Sha LL, Payne AH (1993a). Enzyme characteristics of 2 distinct
forms of mouse 3-beta- hydroxysteroid dehydrogenase/delta-5-delta-4-isomerase
complementary deoxyribonucleic acids expressed in COS-1 cells. Endocrinology

the 21-hydroxylase locus -implications for gene conversion in the human genome.
Nature Genetics 3: 260-265.

Elsevir Science publishing Co., Inc.

specific member of the 3-beta-hydroxysteroid dehydrogenase family, an isoform
possessing an almost exclusive 3- ketosteroid reductase-activity. Journal Of
Biological Chemistry 267: 4513-4517.

DePaoli-Roach AA (1984). Synergistic phosphorylation and activation of ATP-Mg-
dependent phosphoprotein phosphatase by F/A/ GSK-3 and casein kinase II (PC_{0.3}).
Journal of Biological Chemistry 259: 12144-12152.


Dohlman HG, Caron MG, Lefkowitz RJ(1987). A family of receptors coupled to

Doody KM, Carr BR Rainey WE, Byrd W, Murry BA, Strickler RC, Thomas JL,
Mason JI (1990). 3-beta-hydroxysteroid dehydrogenase isomerase in the fetal zone


