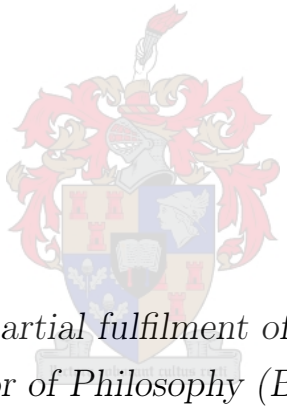
Computational modelling of steroid hormone biosynthesis and metabolism

by

Carla Louw



Thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Biochemistry) in the Faculty of Science at Stellenbosch University

Supervisors: Prof. J.L. Snoep (supervisor)
Dr. D.D. van Niekerk (co-supervisor)

March 2020

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2020

Acknowledgements

I would like to acknowledge the following people and organisations for their contribution:

My supervisor, Prof. Snoep, for his guidance and patience during the completion of this study and giving me the opportunity to truly make this project my own.

My co-supervisor, Dr. Van Niekerk, for his much needed help with the computational aspects of my research, for teaching me valuable skills and for answering countless questions with extraordinary patience.

Prof. P. Swart for his guidance in the field of steroid hormones and Angora goats.

Dr. Y. Engelbrecht for her valuable input during the development of the Angora model.

Prof. K. Storbeck and his lab for a wonderful collaborative journey. A special thank you to Lise Barnard, Monique Barnard and Anna-Mart Burger for helping me understand the experimental procedures behind the data and answering so many questions.

The members of the Molecular Systems Biology Lab for their friendship and encouragement.

Thank you to the following people for data: Erick van Schalkwyk, Dr. R. Conradie, Jonathan Quanson, Monique Barnard, Lise Barnard, Ralie Louw, Dr. Y. Engelbrecht and Prof. K. Storbeck.

The National Research Foundation for funding.

My family and friends, without their love and support the completion of this project would not have been possible.

Contents

Declaration i

Contents iii

List of Figures v

List of Abbreviations / Nomenclature ix

Summary xi

Opsomming xii

1 Introduction 1

2 Literature review: steroid hormones 6
2.1 Mammalian Steroid Hormones 6
2.2 Computational models of steroid biosynthesis 9

3 Literature review: analytical and mathematical methods 15
3.1 Model selection 16
3.2 Linear regression 18
3.3 Nonlinear regression 18
3.4 Regression methods 19
3.5 Identifiability analysis 24
3.6 Confidence intervals 25
3.7 Metabolic control analysis 37
3.8 Computational application: Mathematica package 40

4 A comparative analysis of 17-OHPROG and A4 synthesis in ovine and Angora goat models 43
4.1 Parameterisation and identifiability analysis of the ovine Δ^4 and Δ^5 pathways 44
4.2 Establishing the ovine model as a control model: Comparison with results in literature 58
4.3 Development of the Angora model 70

4.4 Hypocortisolism - comparison of the ovine and Angora models 72

5 Modelling the 5α and 5β reduction of 11-Ketotestosterone in the liver 82

5.1 Experimental data 83
5.2 Model parameterisation 83
5.3 Identifiability analysis 90
5.4 Model validation 90
5.5 Modelling of AKR1D1, SRD5A1 and SRD5A2 kinetics with 1:1:1 enzme expression 91
5.6 Modeling of AKR1D1, SRD5A1 and SRD5A2 kinetics in the liver 94

6 Computational modelling of the intratumoral androgen metabolism in castration resistant prostate cancer 99

6.1 Previously determined parameter values 102
6.2 Newly determined parameter values 103
6.3 Model construction and validation 108
6.4 Identifying possible treatment targets 118
6.5 Increased expression of enzymes 123
6.6 Enzyme inhibition 125

Discussion 128

Appendix A 134

Appendix B 138

Bibliography 144

List of Figures

2.1 Schematic representation of the partial steroid hormone biosynthesis pathway 7

3.1 Theoretical example: Profile likelihood plot of a structurally non-identifiable parameter value . . . 35
3.2 Theoretical example: Data lacking in quantity resulting in practical non-identifiability . . . 36
3.3 Theoretical example: Profile likelihood plot of a practically non-identifiable parameter value . . . 36
3.4 Theoretical example: Data resulting in identifiable parameter value estimations 37
3.5 Theoretical example: Profile likelihood plot of an identifiable parameter value 37

4.1 Schematic representation of the Δ^4 and Δ^5 branches of the steroid hormone biosynthesis pathway . . . 45
4.2 Model fit to the experimental data used for the parameterisation of 3βHSD parameters . . . 50
4.3 Model fit to the experimental data used for the parameterisation of 3βHSD parameters (continued) . . . 51
4.4 Model fit to the experimental data used for the parameterisation of 3βHSD parameters (continued) . . . 52
4.5 Model fit to the experimental data used for the parameterisation of CYP17 parameters . . . 53
4.6 Model validation of the 3βHSD parameters . . . 54
4.7 Model validation of the CYP17 parameters . . . 55
4.8 Identifiability analysis of 3βHSD $Vmax_6$, $Km_{Preg3BHSD}$ and $Km_{Prog3BHSD}$ for the conversion of PREG to PROG . . . 56
4.9 Identifiability analysis of CYP17 $Vmax_5$ and $Km_{17OHProgCYP17}$ for the conversion of 17OH-PROG to A4 . . . 57
4.10 3βHSD and CYP17 activity in Angora goat, Boer goat and Merino sheep adrenal microsomes - figure from literature . . . 59
4.11 Time course data of the PROG metabolism in the Angora goat, Boer goat and Merino sheep - figure from literature . . . 60

4.12 Model simulation of Fig 4.10(a) and 4.10(b) where the multipliers α and β are fitted for . . . 63
4.13 Model simulation of Fig. 4.11 where the multiplier γ is fitted for . . . 65
4.14 Schematic representation of the Δ^4 and Δ^5 branches of the steroid hormone biosynthesis pathway as modelled by Nguyen *et al* . . . 68
4.15 Decreased synthesis of PROG and 17-OHPROG, and increased synthesis of A4 seen with decreased 3βHSD activity. . . . 68
4.16 Changes in steady state concentrations of PROG, A4, DHEA, and 17-OHPROG with varying PREG supply rate. . . . 69
4.17 Time course data and model prediction of the metabolism of 1 μM PREG in the Angora goat. . . . 71
4.18 A comparison of the steady state concentrations of the metabolites in the ovine and Angora models . . . 73
4.19 The effect of increased CYP17 activity relative to 3βHSD activity on the fluxes towards A4 and 17-OHPROG . . . 73
4.20 Partial steroidogenic pathway with numbered branches for MCA . . . 74
4.21 Heat map of the flux control coefficients of the enzymes 3βHSD and CYP17 on the the flux through the Δ4 and Δ5 pathways . . . 75
4.22 Heat map of the concentration control coefficients of the enzymes 3βHSD and CYP17 on the concentrations of metabolites . . . 75
4.23 Effect of increased CYP17 activity relative to 3βHSD activity on the control that the enzymes have on the fluxes J1, J2, J6 and J7 . . . 76
4.24 The difference in the ratio of steady state concentrations of 17-OHPROG and A4 with increased PREG concentration between the ovine model and Angora model. . . . 78
4.25 The differences in the ratio of the fluxes J1 and J6 and the ratio of the fluxes J2 and J7 between the Angora and ovine models with increased PREG concentration . . . 79

5.1 Reactions v1 to v12, as catalysed by either AKR1D1, SRD5A1 or SRD5A2 with A4, T, 11KA4 or 11KT as substrate. . . . 84
5.2 Conversions of T to 5βDHT, 11KT to 11K5βDHT, A4 to 5βDione, and 11KA4 to 11K5βDione by AKR1D1 . . . 87
5.3 Conversions of T to 5αDHT, 11KT to 11K5αDHT, A4 to 5αDione, and 11KA4 to 11K5αDione by SRD5A1 . . . 88
5.4 Conversions of T to 5αDHT, 11KT to 11K5αDHT, A4 to 5αDione, and 11KA4 to 11K5αDione by SRD5A2 . . . 89
5.5 Validation of newly determined AKR1D1, SRD5A1, and SRD5A2 parameter values . . . 91

5.6 Model prediction of the experimental results of 1:1 mixing of AKR1D1 with SRD5A1 . . . 92
5.7 Model prediction of the experimental results of 1:1 mixing of AKR1D1 with SRD5A2 . . . 93
5.8 Activities of AKR1D1, SRD5A1, and SRD5A2 with variation of substrate (11KT and T) concentration . . . 94
5.9 Model simulations of the conversion of 0.8 nM 11KT, T, 11KA4 or A4 to their respective 5α and 5β reduction products at 1:1:1 enzyme expression and physiological enzyme expression levels . . . 95
5.10 Variation in steady state concentrations of 11K5αDHT, 11K5βDHT, 5αDHT, and 5βDHT as the relative expression of AKR1D1 to SRD5A1 and SRD5A2 changes . . . 96
5.11 Model simulation of the steady state concentrations of 11K5αDHT, 11K5βDHT, 5αDHT, and 5βDHT in the liver with physiological expression of AKR1D1, SRD5A1, and SRD5A2 . . . 97

6.1 Schematic representation of the classic androgen and 5αDione biosynthesis pathway . . . 101
6.2 Schematic representation of the 11-oxygenated androgen biosynthesis pathway 101
6.3 Model fit with 95% confidence intervals to experimental 11βHSD2 data . . . 105
6.4 Profile likelihood plots of the 11βHSD2 parameters . . . 107
6.5 Model validation: prediction of 11βHSD2 run-out data . . . 108
6.6 Model construction: fitting of one C42B dataset to determine relative enzyme expression and all 3αHSD mass action parameter values . . . 115
6.7 Model validation: predicting the results of C42B datasets . . . 117
6.8 Model validation: predicting the results of C42B datasets (continued) . . . 117
6.9 Steady state concentrations of classic androgens and 11-oxygenated androgens with a supply of 10 nM 11OHA4 and A4 per hour . . . 118
6.10 Schematic representation of the classic androgen and 5αDione biosynthesis pathway . . . 121
6.11 Schematic representation of the 11-oxygenated androgen biosynthesis pathway 121
6.12 Heat map of the flux control coefficients of the enzymes AKR1C3, 17βHSD2, SRD5A, 11βHSD2, and 3αHSD on the flux through the branches J1 to J28 . 122
6.13 Heat map of the concentration control coefficients of the enzymes AKR1C3, 17βHSD2, SRD5A, 11βHSD2, and 3αHSD on the concentrations of the metabolites T, 5αDione, αDHT, AST, 3αAdiol, 11KA4, 11KT, 11OH5αDione, 11K5αDione, 11K5αDHT, 11OHAST, and 11KAST . . . 122
6.14 Percentage change in steady state concentrations of metabolites with a 100% increase in enzyme expression . . . 124

6.15 Steady state concentrations of metabolites with inhibition of AKR1C3 and 17βHSD2 125
6.16 Steady state concentrations of metabolites with inhibition of SRD5A, 11βHSD2, and 3αHSD 126

List of Abbreviations / Nomenclature

3αHSD	3α-Hydroxysteroid dehydrogenase
3βHSD	3β-Hydroxysteroid dehydrogenase
5αDHT	5α-Dihydrotestosterone
5βDHT	5β-Dihydrotestosterone
11βHSD2	11β-Hydroxysteroid dehydrogenase 2
11K5αDHT	11-Keto-5α-Dihydrotestosterone
11K5βDHT	11-Keto-5β-Dihydrotestosterone
11K5αDione	11-Keto-5α-Dione
11K5βDione	11-Keto-5β-Dione
11KA4	11-Ketoandrostenedione
11KAST	11-Ketoandrosterone
11KT	11-Ketotestosterone
11OH5αDione	11-Hydroxy-5α-Dione
11OHA4	11β-Hydroxyandrostenedione
11OHAST	11β-Hydroxyandrosterone
16-OHPROG	16-Hydroxyprogesterone
17βHSD	17β-Hydroxysteroid dehydrogenase
17βHSD2	17β-Hydroxysteroid dehydrogenase 2
17-OHPREG	17-Hydroxypregnenolone
17-OHPROG	17-Hydroxyprogesterone
A4	Androstenedione
AKR1C3	Aldo-keto reductase 1C3
AKR1D1	Aldo-keto reductase 1D1
CAH	Congenital adrenal hyperplasia
CI	Confidence intervals
CORT	Corticosterone
CRPC	Castration resistant prostate cancer
CYP11A1	Cytochrome P450 11A1
CYP11B1	Cytochrome P450 11B1
CYP11B2	Cytochrome P450 11B2

CYP17	Cytochrome P450 17α-Hydroxylase/17,20-Lyase
CYP21	Cytochrome P450 21 or Steroid 21-Hydroxylase
DHEA	Dehydroepiandrosterone
DHT	Dihydrotestosterone
DOC	Deoxycorticosterone
DOCL	Deoxycortisol
E1	Estrone
E2	Estradiol
EACs	Endocrine active chemicals
EDCs	Endocrine disrupting chemicals
ER	Estrogen receptor
MCA	Metabolic control analysis
MET	Metyrapone
NAD^+	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
ODE	Ordinary differential equation
PCOS	Polycystic ovary syndrome
PREG	Pregnenolone
PROG	Progesterone
SBP	Steroid binding protein
SRD5A	Steroid 5α-reductase
SRD5A1	Steroid 5α-reductase 1
SRD5A2	Steroid 5α-reductase 2
SSR	Sum squared residual
T	Testosterone
ZF	Zona fasciculata
ZG	Zona glomerulosa
ZR	Zona reticularis

Summary

This study describes the use of computational modelling and statistical techniques to address three topics in the field of steroid hormone research. The first is the hypocortisolism seen in the South African Angora goat. In a comparative analysis the construction, parameterisation and validation of a model describing the Δ^4 and Δ^5 pathways in both the ovine and Angora goat species are completed. With these models the issue of identifying a possible treatment target is addressed. The second topic is the steroidogenic activity in the human liver. This includes the construction, parameterisation and validation of a model describing the relative conversion of classic androgens and 11-oxygenated androgens to their respective 5α or 5β reduction products in the human liver. The model indicates that under physiological steady state conditions the 5β reduction of both the classic androgens and the 11-oxygenated androgens are the preferred reaction. The third and final topic discussed in this study is the steroidogenic activity in prostate cancer C42B cells. The construction, parameterisation and validation of a model describing the steroid hormone biosynthesis in the C42B castration resistant prostate cancer cell line is included. Three possible treatment targets for castration resistant prostate cancer in C42B cells are identified. This study also describes the development of an add-on Mathematica package, `IdentifiabilityAnalysis`, which simplifies the process of model parameterisation and identifiability analysis. This package is used throughout this study.

Opsomming

Hierdie studie beskryf die gebruik van wiskundige modellering en statistiese tegnieke om drie onderwerpe op die veld van steroïedhormoon navorsing aan te spreek. Die eerste is die hipokortisolisme wat in die Suid-Afrikaanse Angorabok gesien word. In 'n vergelykende analise is die konstruksie, parameterisering en validering van 'n model wat die Δ^4 and Δ^5 padweë beskryf, beide in die skaap- en Angorabokspesies voltooi. Met hierdie modelle word die kwessie van die identifisering van 'n moontlike teiken vir behandeling aangespreek. Die tweede onderwerp is die steroïdogeniese aktiwiteit in die lewer van die mens. Dit sluit in die konstruksie, parameterisering en validering van 'n model wat die relatiewe omskakeling van klassieke androgene en 11-geöksideerde androgene na hul onderskeie 5α of 5β reduksieprodukte in die menslike lewer beskryf. Die model dui aan dat die 5β reduksie van beide die klassieke androgene en die 11-geöksideerde androgene by fisiologiese bestendige toestand die voorkeurreaksie is. Die derde en laaste onderwerp wat in hierdie studie bespreek word, is die steroïdogeniese aktiwiteit in prostaatkanker C42B-selle. Die konstruksie, parameterisering en validering van 'n model wat die biosintese van die steroïedhormone in die C42B-kastrasie-weerstandige prostaatkanker-sellyn beskryf, is ingesluit. Drie moontlike teikens vir die behandeling van kastrasie-weerstandige prostaatkanker in C42B-selle word geïdentifiseer. Hierdie studie beskryf ook die ontwikkeling van 'n addisionele Mathematica-pakket, `IdentifiabilityAnalysis`, wat die proses van modelparameterisering en identifiseerbaarheidsanalise vergemaklik. Hierdie pakket word dwarsdeur hierdie studie gebruik.

Chapter 1

Introduction

Steroid hormones play a key role in various processes in the body during all stages of life and are classified into three main groups. Glucocorticoids regulate the immune system and inflammatory responses, while mineralocorticoids maintain a healthy balance of sodium and water in the body. The third group, sex steroids, regulate sexual development and growth. These steroid hormones are all synthesised from a common precursor, cholesterol, by an array of different enzymes. The cytochrome P450 enzymes and the hydroxysteroid dehydrogenase enzymes are the two main groups into which these enzymes fall. Many of these enzymes are multi-functional and are expressed in different tissues in the body. As such, these enzymes form intricate steroidogenic systems throughout the body. Imbalances and defects in any of these steroidogenic pathways can lead to a number of different diseases [70, 86, 102].

The main aim of this study is to demonstrate the application of analytical techniques to construct, parameterise, and validate models of different steroidogenic systems. Computational modelling of these complex systems has aided the study of steroid hormone biosynthesis in various species. Such models include the model by Becker *et al*, which was the first model describing testicular steroidogenesis [9]. Breen *et al* [15] used mathematical modelling of the intraovarian metabolic network in fish to study the effects of endocrine active compounds in the environment on hormone synthesis. Murphy *et al* [76] conducted a similar study by modelling the effects of endocrine disrupting chemicals on two species of fish. Another study of the effects of endocrine disrupting compounds on steroidogenesis was completed by Saito *et al* [93], who studied these effects on the human endocrine system. Breen *et al* [13] also completed their own independent study of these effects on human H295R cells with the use of computational modelling. Nguyen *et al* [79–81] conducted extensive studies on steroid hormone synthesis in mammals and constructed an accompanying model describing the dynamics of the Δ^4 and the Δ^5 steroidogenic pathways. Cook *et al* [22] modelled the oxidation of androsterone and androstenedione in humans to study castration resistant prostate cancer. Another study modelling hormone synthesis in humans was conducted by Selgrade and Schlosser [96, 97]

who studied the nine stages of hormone synthesis in the ovaries. Chapter 2 contains an in depth literature review of the examples of steroidogenic modelling mentioned above, although the modelling of steroid hormone synthesis is not limited to these examples.

A comparative analysis of 17-OHPROG and A4 synthesis in ovine and Angora goat models

Computational models have aided in the understanding of steroid hormone biosynthesis. One question, which remains unanswered, is the dynamics of hypocortisolism in the South African Angora goats. The mohair industry in South Africa suffers the loss of young Angora goats during cold spells as these animals are not able to produce sufficient levels of cortisol to survive physiological stress. With South Africa being the leading producer of mohair worldwide, this problem has great financial implications [74]. It was found that the Angora goat has increased CYP17 activity, relative to 3βHSD activity [28]. These two enzymes catalyse the reactions of the Δ^4 and Δ^5 pathways in steroid biosynthesis, eventually leading to the synthesis of either aldosterone, cortisol or androstenedione [26–28, 118]. A comparative study by Engelbrecht *et al* [28] between the Angora goat and another wool producing, but more hardy livestock species, the Merino sheep, determined that there is little difference in the synthesis of aldosterone between the animals, but a greater difference in the synthesis of cortisol and androstenedione [26–28].

Due to the multi-functionality of the two enzymes of the Δ^4 and Δ^5 pathways, it is difficult to pinpoint exactly how the increased CYP17 activity affects the flux of steroids through the pathway. It is not known which of the two enzymes has more control over the flux and the concentrations of the system, and where exactly in the pathway the most control lies. The effect of a stress response on the flux through the pathways an how it differs from the system without stress has also not been investigated. Once this is known for the Angora goat and ovine species a comparison can be made to identify the origin of hypocortisolism and finally some possible treatment target or strategy can be identified.

In chapter 4 we address these questions of hypocortisolism in the Angora goats with the aid of computational modelling. The first aim is to construct and validate a model for steroid hormone synthesis in the ovine species, which is used as a control model for steroid biosynthesis. This model will describe the Δ^4 and the Δ^5 branches of the steroidogenic pathway, leading to the synthesis of either cortisol or androstenedione. The first objective of this aim is to parameterise the model with experimental transfection data. The second objective is to validate the ovine model against experimental transfection data as well as animal data from literature. The third objective is to validate this model against the theoretical model published by Nguyen *et al* [81]. The second aim is to create a model describing these same Δ^4 and Δ^5 branches of the pathway, but for the Angora goat.

The first objective is to construct the Angora model by adapting the ovine model. The second objective is to validate the Angora model against experimental transfection data in literature. The third aim is to analyse and compare the dynamics of the ovine and the Angora model. The objective is to conduct metabolic control analysis and steady state analysis of the models. This will show how the multi-functional dynamics of the enzymes of the Δ^4 and Δ^5 pathways lead to the hypocortisolism in the South African Angora goats.

Modelling the 5α and 5β reduction of 11-Ketotestosterone in the liver

Computational models can also aid in the process of understanding and describing the dynamics of novel steroidogenic pathways. In recent years the role of 11-Ketotestosterone has been proven as a potent androgen [87, 108]. Studies are also being undertaken to determine the involvement of 11-oxygenated androgens in various hormone sensitive diseases such as polycystic ovary syndrome, congenial adrenal hyperplasia, and castration resistant prostate cancer [7, 77, 87]. The 11-oxygenated androgens are equipotent to the classic androgens, however, these androgens have not yet been studied in the same amount of detail as the classic androgens. As such there is still much to be learnt of their dynamics, especially when pooled with the classic androgens. Both the classic androgens and the 11-oxygenatated androgens can be converted to different products, depending on which enzyme it binds to. In the liver, where the enzymes AKR1D1, SRD5A1, and SRD5A2 are expressed, these androgens can be either 5α reduced by SRD5A1 or SRD5A2 or 5β reduced by AKR1D1. The 5α products are potent androgens, but the 5β products are not. As these enzymes have different efficiencies for the different androgen substrates and are present in the liver at different expression levels, it is difficult to predict how much of each androgen substrate will be converted to potent 5α products or non-potent the 5β products.

The main research aim of Chapter 5 is to predict ratios of these 5α and 5β products of both the classic and 11-oxygenated androgens. The first objective is to construct and parameteris a model of the AKR1D1, SRD5A1, and SRD5A2 enzyme kinetics in the human liver is with experimental transfection data. The second objective is to include the physiological expression levels of these enzymes in the model, whereafter the third objective is to study the model simulations at steady state. We are able to determine the relative conversion of testosterone and 11-Ketotestosterone to their respective 5α and 5β reduction products in the liver of both men and women.

Computational modelling of the intratumoral androgen metabolism in castration resistant prostate cancer

Modelling of steroid hormone biosynthesis can also be used to identify possible treatment targets for hormone sensitive diseases, such as castration resistant prostate cancer. This type of cancer often emerges after the initial treatment of chemical or physical castration for prostate cancer fails. The pathway that is followed in the synthesis of potent androgens in castration resistant prostate cancer is the 5αDione pathway. It has also been found that the 11-oxygenated androgens can serve as substrates for this type of cancer [7, 45, 98]. Similar to the systems mentioned above, the enzymes catalysing these reactions are multi-functional. As such, with the classic androgens and 11-oxygenated androgens pooled, these enzymes can convert androgens from either group into their respective intermediates and products. Knowing how these enzymes interact when competing for the same substrates will aid in the search of novel treatment targets for castration resistant prostate cancer.

In chapter 6 we aim to identify possible treatment targets for castration resistant prostate cancer in C42B cancer cells with the aid of mathematical modelling. The first objective is to construct a model that describes the combined dynamics of the 5αDione and the 11-oxygenated androgen pathway in C42B cancer cells. The second objective is to study the model at physiological steady state and the effects of inhibiting and up-regulating the enzymes. We identify three possible treatment targets for castration resistant prostate cancer in C42B cancer cells.

Chapter 3 is a review of analytical techniques and describes the development of a novel Mathematica add-on package `IdentifiabilityAnalysis`. The functions of this package have been used throughout this project to analyse experimental data and construct the accompanying steroidogenic models. The use of the package is however not limited to steroid hormone research. The package contains five functions which include fitting of the user input data to a model described by ordinary differential equations, testing the identifiability of newly determined parameter values, as well as creating profile likelihood plots. The aim in creating this package was to simplify the process of model parameterisation and identifiability analysis, by minimising the amount of coding needed to be completed by the user.

In summary, this study consists of two review chapters and three research chapters. Chapter 2 is the first review chapter and gives a brief overview of mammalian steroid hormones and a literature review of mathematical models of steroid biosynthesis in both humans and other animals. Chapter 3, the second review chapter, is a summary of the analytical methods used during data analysis and model construction. This chapter also includes an outline of the `IdentifiabilityAnalysis` package. Chapter 4 is the first

research chapter and contains the development of the models describing the Δ^4 and Δ^5 pathways in the ovine and Angora goat species. The second research chapter, chapter 5, outlines the development of the model describing the 5α and 5β reduction of classic androgens and 11-oxygenated androgens in the human liver. Chapter 6 is the final research chapter in this study and includes the development of the model describing the androgen metabolism in castration resistant prostate cancer C42B cells. This study is concluded with a general discussion, Appendix A which contains the identifiability analysis results of chapter 4, and Appendix B which contains the identifiability analysis results of chapter 5.

Chapter 2

Literature review: steroid hormones

This chapter contains a brief overview of mammalian steroid hormones, the zonation of adrenal cortex and diseases caused by any abnormalities or deficiencies in the steroidogenic pathways. A literature review of mathematical models describing steroidogenesis in various species is included. A review of two models in literature, relating to the work done in this study, concludes this chapter.

2.1 Mammalian Steroid Hormones

Three classes of steroids are synthesised and secreted in mammals by the adrenal cortex; the mineralocorticoids, the glucocorticoids and the gonadocorticoids [21, 102]. These steroids have various physiological functions throughout the body [15]. Figure 2.1 shows the partial steroidogenic pathway. Cholesterol, a steroid hormone precursor, is converted into steroid hormones by the cytochrome P450 and hydroxysteroid dehydrogenase enzymes [69, 70, 79, 86, 102]. These steroid hormones are 3βHSD, CYP11A1, CYP11B1, CYP11B2, CYP17, and CYP21 [21, 69, 70]. Cholesterol is converted to Pregnenolone (PREG) by CYP11A1 whereafter either 3βHSD or CYP17 converts PREG to Progesterone (PROG) or 17-hydroxypregnenolone (17-OHPREG) respectively. CYP17 can also convert 17-OHPREG to dehydroepiandrosterone (DHEA) as well as convert PROG to 17-hydroxyprogesterone (17-OHPROG) and in turn convert 17-OHPROG to androstenedione (A4). The conversion of 17-OHPREG to DHEA by the lyase activity of the CYP17 enzyme is more effective than that of the conversion of 17-OHPROG to A4, specifically in the human, primate, ovine and bovine species. In these species A4 is mainly produced from DHEA. CYP17 also synthesises 16-hydroxyprogesterone (16-OHPROG) from PROG along with 17-OHPROG. The synthesis of 16-OHPROG is more prominent in humans and primates, while synthesis of 16-OHPROG in goats and sheep are present, but minimal [21, 47, 69, 70, 79, 86, 110]. The hormones 17-OHPREG and DHEA can be converted to 17-OHPROG and A4 by 3βHSD respectively [21]. A4 is then converted to

Testosterone (T) by 17βHSD [102]. CYP21 can then convert PROG and 17-OHPROG to deoxycorticosterone (DOC) and deoxycortisol (DOCL) respectively. DOC and DOCL are in turn converted to corticosterone (CORT) and cortisol by CYP11B1 respectively. CYP11B2 then finally converts CORT to 18-hydroxycorticosterone and that in turn to aldosterone [70, 86, 102]. In the species of human, sheep, goat and cattle, the Δ^5 pathway is preferred for the synthesis of A4 [21, 70, 103, 111]. In both the porcine and rodent species A4 is synthesised via either the Δ^4 or Δ^5 pathways, the Δ^4 pathway is preferred in rodents [6, 21, 32, 63].

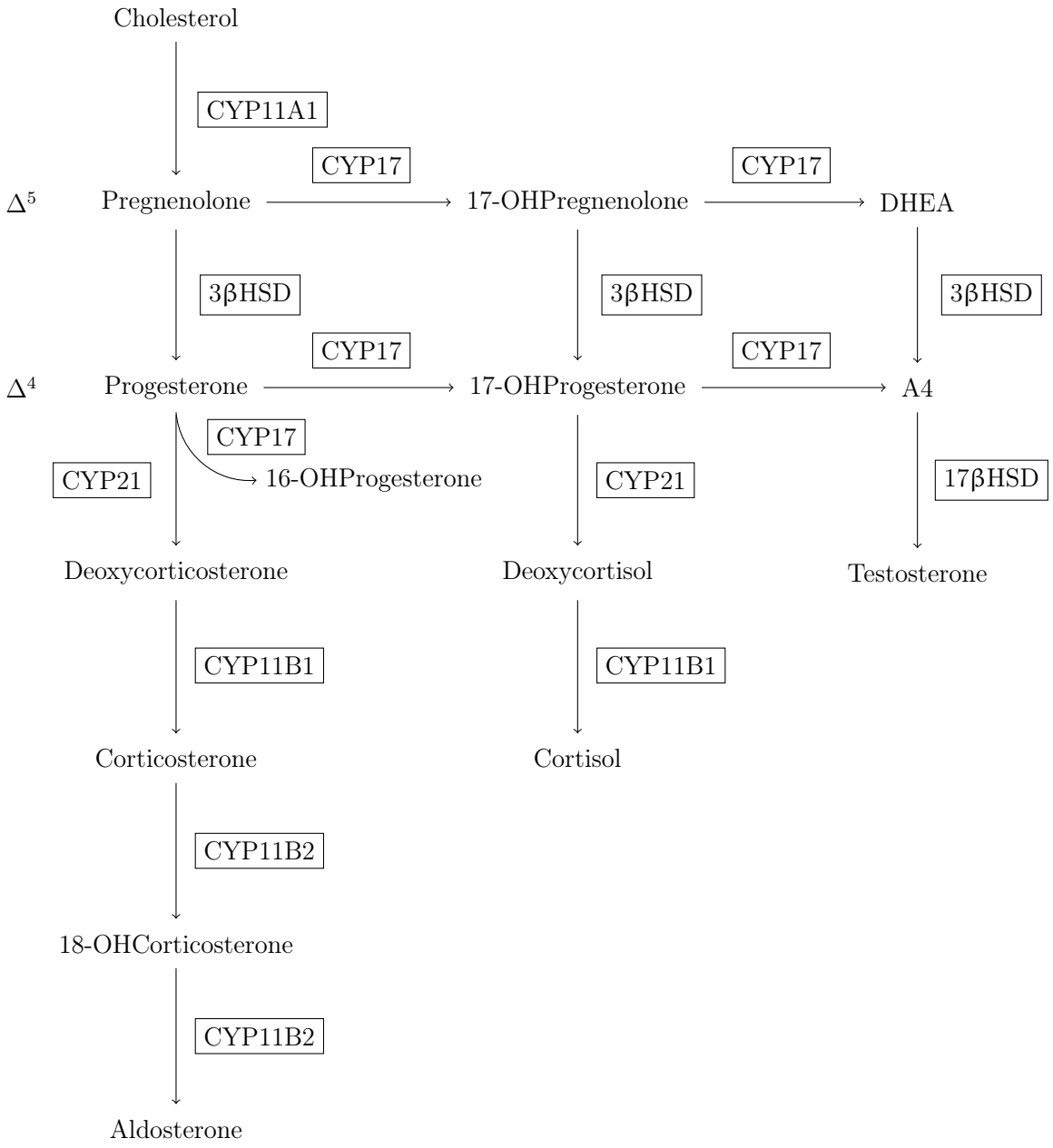


Figure 2.1: **Schematic representation of the partial steroid hormone biosynthesis pathway.**

Steroid hormones are synthesised on demand as cells do not store these hormones. The levels of steroid substrates and the steroidogenic enzyme activity regulate steroid hormone biosynthesis [91].

2.1.1 Zonation of the adrenal cortex

The adrenal cortex consists of three zones; the outer zona glomerulosa (ZG), the inner zona fasciculata (ZF), and the zona reticularis (ZR). The ZF makes up the majority of the adrenal cortex (about 75%), while the ZG only makes up about 15% of the adrenal cortex, with the remaining 10% consisting of the ZR. Not all steroidogenic enzymes are expressed in each of the three zones and as a result the synthesis of steroid hormones are zone specific. [21, 70, 86, 102].

2.1.2 Diseases

There are various diseases caused by imbalances in steroid hormone biosynthesis. Congenial adrenal hyperplasia (CAH) can be caused by a deficiency in either 3βHSD or CYP17, however CYP21A2 deficiency is most common [84, 85, 102]. This leads to decreased production of cortisol, aldosterone, PROG, androgens, and estrogens [69, 70, 84, 99, 102]. Disruption in normal 3βHSD activity can have a critical effect on physiological processes, This is caused by disruptions in the normal production of the steroid hormones [70, 72, 84, 99]. Polycystic ovary syndrome (PCOS) leads to increased androgen secretion, caused by an irregularity in the steroidogenic pathway [69, 90, 116]. Cushing Syndrome is caused by increased levels of glucocorticoids in the system. Symptoms include obesity, acne, hirsutism in female patients, depression, lethargy, insomnia, and amongst others, hypertension [102]. Patients can also suffer from glucocorticoid deficiency. This includes Addison Disease, Autoimmune Adrenalitis and Hypoadrenalism [102]. Hormone dependent cancer, such as breast cancer and prostate cancer, are also influenced by the dynamics of steroid hormone biosynthesis [1, 7, 71].

A comprehensive understanding of steroid biosynthesis is fundamental for identifying possible branches in the pathway to target for treatment of endocrinological disorders to ensure proper steroid synthesis [21, 65, 79]. Imbalances in the steroidogeneic pathway can have a negative effect on normal physiological functions. It is however not entirely clear how these imbalances influence steroid biosynthesis exactly [38]. Many studies have been conducted with the aim of understanding the differences in flux through the pathways in different species [21, 31, 79, 80]. Steroidogenic models promote the understanding of biochemical processes and responses to stimuli and could aid in identifying predictive biomarkers. If properly constructed such a model could simulate the therapeutic or negative effects that compounds may have on hormone biosynthesis [15, 38].

2.2 Computational models of steroid biosynthesis

To gain insight into steroid biosynthesis, studying the differences in hormone production in different species can be beneficial. Differences and similarities in enzymatic activity between species can grant insight into steroid biosynthesis at a biochemical and physiological level [21, 27, 28]. Mathematical models that accompany the experimental studies aid the process of understanding the flux of steroids through biochemical pathways [15, 79–82, 125]. It is often challenging to replicate the steroidogenic pathways experimentally, especially more complicated pathways where the inhibition of one branch may affect the other parts of the pathway. Another limitation is the financial implications with the completion of a large number of experimental procedures to gain understanding of steroid biosynthesis [22]. Mathematical models are not subject to such limitations [22, 82]. With mathematical models existing knowledge can be combined, making it easier to identify gaps in the current research [22]. Although *in silico* analysis of a system does not necessarily eliminate the need for experimental work, it can aid in the process of experimental design [22].

In 1980 Becker *et al* [9] created the first mathematical model that describes testicular steroidogenesis. The model was created with two goals in mind: to predict the secretion rates of steroids during biosynthesis, and to test the validity of assumptions used in studies of testicular T synthesis. This model is specific to rabbits and rats. Model validation consisted of comparing predicted steroid secretion rates with those of *in vitro* perfused control testes. The predominant pathway for T synthesis from the PREG precursor in rabbit and rats testes was also determined through analyses.

Breen [15] created a steady state mechanistic model of the intraovarian metabolic network. The model describes the synthesis of T and estradiol (E2) as well as the influence of endocrine active compounds of therapeutic or environmental nature on the hormone synthesis. The dose-response behaviour of endocrine active compounds were studied and the model was developed to show the biochemical response induced by perturbations at a physiological level. The model was parameterised for the Cyprinidae species (fathead minnow). The ecological significance of this small fish makes it a good animal model. In addition the Cyprinidae species is a large family of fish that has been studied in great detail. Much is know about culturing of the fish and its life cycle, as well as how to manipulate its reproductive cycle. Its 4-5 month life cycle and sufficient ovary tissue volume at maturation that can be used for steroidogenic assays also contribute to its appeal as an animal model. The design of the model includes the metabolic pathway and the transport pathway. It describes the enzyme kinetics in both the culture medium and the ovary tissue where the transport pathway includes the uptake of substrate (cholesterol) and the secretion of products (A4, estrone, T and E2). The metabolic

pathway includes the conversion of substrate to products and the enzyme inhibition by Fadrozone, an endocrine-active chemical (EAC) which results in competitive enzyme inhibition. Parameter values were determined by a least squares fitting of the model to mean concentration values from replicate studies. Two parameter values were taken from literature while the remaining 11 parameters were fitted for. V_{max} and K_m values for each branch in the pathway were fitted for as a single rate constant. A sensitivity analysis of the 11 parameters showed which parameters resulted in the largest changes in model output induced by small changes in the parameter values. The model is a good example of the use of a steroidogenic model to predict *in vitro* assay results.

EACs can alter normal human endocrine function. Breen *et al* [13] developed a computational model describing the effect of EACs on steroidogenesis. The model was parameterised with *in vitro* data of steroidogenesis in H295R cells. The model includes the culture medium and the H295R cells as two compartments. The transport pathway of cholesterol uptake forms part of the model dynamics as well as the metabolic pathways of hormone metabolism and secretion. The metabolism of the following hormones are included: PREG, 17-OHPREG, DHEA, PROG, 17-OHPROG, A4, T, DOC, CORT, aldosterone, DOCL, cortisol, estrone (E1), and E2. The enzyme inhibitory effects of metyrapone (MET) also forms part of the model dynamics. MET is and EAC that inhibits the enzyme CYP11B1. Sensitivity analysis showed that the 17-OHPREG pathway is preferred over the PROG pathway for the synthesis of CORT [14]. This model was later expanded by Breen *et al* to include cell proliferation and the synthesis of oxysterols.

Murphy *et al* [76] developed a model that describes the synthesis of vitellogenin in mature female sciaenid fish. Two model-organisms of the Scianidae family is used, spotted seatrout (*Cynoscion nebulosus*) and Atlantic croaker (*Micropogonias undulatus*). Both of these animals are well-established models for studying reproductive endocrinology and endocrine toxicity. Vitellogenin is a yolk precursor protein and proper synthesis thereof during the larval fish life stage is crucial. Chemical compounds and environmental factors can alter the reproduction of fish by disrupting endocrine function. High concentrations of trace elements, like cadmium, can hinder the processes of gonadotropin regulation and steroidogenesis. Gonadotropin and vitellogenin are biomarkers for disruptions in endocrine and reproductive function. The model describes the effect of two endocrine disrupting chemicals (EDCs) on vitellogenin synthesis in fish. The model consists of eight ODEs showing vitellogenin synthesis in a mature female fish over a period of six months with hourly introduction of gonadotropin. The ODEs correspond to the rate of change of the following metabolites: T, steroid binding protein (SBP), SBP bound to T, E2, SBP bound to E2, estrogen receptor (ER), ER bound to E2, and vitellogenin. The model is used to simulate the disruptive effects of EDCs on steroidogenesis and

vitellogenisis and is compared to field data.

Saito *et al* [93] created a model, describing steroidogenesis in humans, that allowed them to study the effects of various endocrine disrupting compounds on human endocrine function. Human adrenocortical carcinoma NCI-H295R cells were used for collecting *in vitro* experimental data of steroid hormone synthesis. These results were used to parameterise the mathematical model, along with parameter values taken form literature. The model includes four important processes: cell proliferation, intracellular cholesterol translocation, the diffusional transport of hormones, and the metabolic pathways of adrenal hormone synthesis. This comprehensive model includes 14 steroids in the steroidogenic pathway (starting with cholesterol), metabolised by nine enzymes. Results of sensitivity analysis suggested that the enzyme 3βHSD has the greatest influence on steroid biosynthesis. Saito *et al* also monitored which enzymes were most affected by adrenal toxic substances (11 compounds were tested). Adrenal toxicity is an unavoidable concern during the drug development process.

Selgrade and Schlosser [96, 97] developed a model describing the synthesis of the ovarian hormones E2, PROG, and inhibin, and the gonadotropin hormones, luteinizing hormone and follicle stimulating hormone. These five hormones play a role in the regulation and maintenance of a normal menstrual cycle. The model was created with the goal in mind to test he effects of exogenous compounds on the human menstrual cycle, therefore a model of a normal menstrual cycle was developed. The model is a system of linear ODEs describing the blood levels of these hormones during the nine stages of hormone synthesis in the ovaries. Model validation was done by comparing the model to data in literature of women with normal menstrual cycles.

The following work by Cook *et al* [22] and Nguyen *et al* [79, 81] relates most to the work shown in this study.

Cook *et al* [22] developed a model of the oxidation of androsterone to androstanedione in humans. The model was developed to aid in the understanding of androgen synthesis and the progression of castration resistant prostate cancer (CRPC). Prostate cancer cell proliferation is triggered by the androgens T and dihydrotestosterone (DHT). Treatment is administered in the form of surgical or chemical castration to reduce the levels of circulating T. DHT can be synthesised via three pathways, the frontdoor pathway and two backdoor pathways. DHT is synthesised from T via the frontdoor pathway, however, the two backdoor pathways do not require T as an intermediate metabolite in the production of DHT. This model only describes one branch in the pathway. The model was parameterised with values from from literature and with experimental data.

These classic androgen pathways, as discussed in the above mentioned study by Cook *et al* is reparameterised for C42B CRPC cells in chapter 6. The dynamics of the 11-oxygenated androgen pathway is also included to consider the effects of both the classic androgens and the 11-oxygenated androgens on the dynamics of CRPC in C42B cells.

Nguyen [79, 81] investigated the effect that a change in the PREG supply can have on the synthesis of steroid hormones. The paradoxical increase of E2 in rhesus monkey and sheep with inhibition of 3βHSD is studied. The synthesis of androgens and estrogens in the species human, primate, ovine and bovine is studied with a mathematical model. The model describes minimal conversion of 17-OHPROG to A4 with the majority of A4 being synthesised via the Δ^5 pathway. The model includes a constant supply of PREG and constant secretion of all steroid intermediates and products, including PREG. The steroid secretion rates are dependent on the steroid concentration and a coefficient value, used for all steroids in the pathway. The maximum reaction rates for an enzyme is assumed to be the same for all reactions catalysed by the enzyme. With this model it is shown that the three main influences on the rate of steroid production are the availability of substrate, the rate at which steroid hormones are secreted, as well as the enzyme activities. The specific changes in the rate of steroid biosynthesis caused by these three influences is dependent on the biochemical nature of the enzyme reactions. The model includes the competitive inhibition of the different substrates that are metabolised by the same enzymes. The effect that varying enzyme activities and substrate levels can have on the synthesis of the final products (estrogens and androgens) is shown. This model was constructed with parameter values taken from various studies in literature. It is a theoretical model which describes the dynamics of the Δ^4 and Δ^5 pathways in general in the human, non-human primate, ovine and bovine species. This model also does not take into account the synthesis of 16-OHPROG, which is not synthesised in negligible levels in human and non-human primate species [110].

Steroidogenic enzymes are compartmentalised and many are membrane bound. Both CYP17 and 3βHSD are found in the endoplasmic reticulum, however 3βHSD is also found in the mitochondria. As steroids have high intracellular diffusion coefficients the compartmentalisation should not have a great effect on steroid synthesis. Nguyen investigated the effect that the compartmentalisation of enzymes could have on the rate of steroid synthesis [80]. A mathematical model of the reaction-diffusion kinetics of CYP17 and 3βHSD was used. It was determined that the rate of steroid synthesis is not greatly influenced by the spatial separation of the enzymes within the endoplasmic reticulum. The rate of steroid synthesis is however influenced by the separation of enzymes in different cells. When the enzymes CYP17 and 3βHSD are in different cells, an increased distance between the cells leads to decreased steroid production. This

separation of CYP17 and 3βHSD increases the ratio between the Δ^5 pathway and the Δ^4 pathway when CYP17 and the PREG supply are in the same cell. It was determined that no compartmentalisation of any kind had any effect on the qualitative result of steroid synthesis with variation in enzyme activity or PREG supply. Although the absolute concentrations of the synthesised steroids changed, the levels of the steroids relative to each other remained fairly similar [80].

This general model for steroidogenesis, as developed by Nguyen *et al* [81], is adapted and parameterised for the ovine species as well as the South African Angora goat in chapter 4.

The South African Angora goat is an example of an experimental system showing the effects of altered enzyme activity levels on steroid biosynthesis. The Angora goat suffers from reduced cortisol production (hypocortisolism) [27, 28, 118]. The Boer goat and Merino sheep are considered to be good control models as they do not experience this reduced cortisol production seen in the Angora goats [27, 28].

Many studies have been conducted toward understand the hypocortisolism in Angora goats [27, 28, 39, 48, 103, 104, 106, 118, 123]. It was found that both CYP17 and 3βHSD contribute to the hypocortisolism seen in the Angora goats. Identifying the cause of the hypocortisolism proved to be complicated as there is more than one factor contributing to the imbalances in the steroidogenesis [27, 28, 38, 39].

Whether cortisol is synthesised via the Δ^4 or Δ^5 pathway is determined by the expression levels of 3βHSD and CYP17 relative to each other [10, 21, 111]. CYP17 and 3βHSD compete for the same steroid substrates (PREG and 17-OHPREG). The relative expression of these two enzymes, their substrate specificities as well as their activities all influence the flux of steroid hormones through the pathway [10, 21, 38, 70, 79, 81, 111]. A lower expression of CYP17 relative to 3βHSD will favour the synthesis of glucocorticoids while a higher expression of CYP17 relative to 3βHSD will favour the synthesis of androgens [10, 21]. The 3βHSD enzyme's activity is subject to product inhibition [79]. The reactions of 3βHSD are reversible while cytochrome P450 reactions are irreversible [70]. Due to this complexity it is challenging to predict the effect that flux changes through the branches of the pathway might have on the synthesis of steroids [79–81]. These steroidogenic processes are complex and the experiments are expensive. The use of models that accurately describe the hormone biosynthesis for both a control species and the steroid deficient species could help circumvent these limitations.

Dr R. Conradie (at Stellenbosch University, unpublished) created a computational model of the Δ^5 steroidogenic reactions of the CYP17 enzyme. The model showed the change in concentrations of PREG, 17-OHPREG and DHEA over time.

Van Schalkwyk [119] expanded on this model by including the Δ^4 reactions of CYP17 as well as the reactions catalysed by 3βHSD. Van Schalkwyk studied the kinetics of both CYP17 and 3βHSD and the effects that they have on each other. A computational model was constructed of the partial steroidogenic pathway consisting of these two enzymes. The model describes the change in concentration of the substrates, intermediates and products of the pathway over time. Michaelis-Menten equations with product inhibition were used to describe the enzyme kinetics. The rate equations included the enzymes CYP17 and 3βHSD competing for the same substrates and intermediates. These rate equations form the ordinary differential equations used to construct the model which was used to estimate the kinetic parameters. The parameters for the reactions catalysed by CYP17 and 3βHSD were estimated separately for the two enzymes. Progress curve data for the conversion of PREG to the respective intermediated and products by CYP17 and 3βHSD were used. Initial rate analysis was used to normalise these experiments. The parameter values subsequently obtained were used to construct the model. This model describes the competitive nature of the two enzymes as well as the steroid flux through the pathway. Although the model was validated, no identifiability analyses were done on the parameter estimations. The following chapter is a review of the analytical methods used throughout the rest of this study.

Chapter 3

Literature review: analytical and mathematical methods

This chapter is an overview of some of the analytical methods used during the construction of ordinary differential equation (ODE) based models. The aim of many scientific studies is to solve some form of the inverse problem. The inverse problem is to determine the structure and functional dynamics of a system from experimental observations. Mathematical model selection based on the known information about the system is the first step towards solving the problem. Experimental design (subject to resource availability, and financial and time constraints) is the next important step. As the number of experiments that can be completed is often limited, much importance lies in designing the experiments so that the results will contain as much of the needed information as possible. The experiments are then conducted and the results analysed, whereafter the final step in solving the inverse problem is completing identifiability analysis [19, 30, 55, 126].

Mathematical models are used for analysing data, drawing conclusions from data and predicting future outcomes based on current data [83, 126]. Regression analysis can help with the development of accurate models, and is used for studying the relationship between variables, describing data, predicting outcomes and estimating parameter values. This statistical technique is used in multiple fields of research, including the field of biochemistry for the study of enzyme catalytic reactions [73, 126].

For detailed mechanistic models of enzyme catalytic reactions, the process of data fitting can be divided into the following steps: the first step is to determine a suitable mechanism that describes the enzyme kinetics. The model is then fitted to the experimental data by means of linear or non-linear regression [17, 18, 55, 126]. An objective function is used to minimise the difference between the observed data and the model fit. The method by which the parameter values are solved for numerically needs to be specified. Thereafter the model is validated to determine whether the model is an appropriate description of

the data. Part of this validation step is identifiability analysis. If the model is shown not to be appropriate for the data, this process is repeated [58, 61].

This chapter serves as an overview of some of the methods and techniques available that can be used during each of the steps mentioned above. Model selection is discussed first, followed by linear and nonlinear regression methods. Identifiability analysis of parameters, and methods of determining their confidence intervals are discussed, and finally metabolic control analysis is briefly described.

3.1 Model selection

There are various classic mathematical methods that have been used for decades to develop models for biological systems. These include ordinary regression models, generalised linear and nonlinear regression models and growth curve models [83]. The model type as well as the number of parameters will determine which method is to be used [5].

Model selection should not only depend on the dynamics of the system, but the type and quantity of the available data should also be considered. The enzyme kinetics of the biological systems addressed in this study can be described with reversible or irreversible Michaelis-Menten kinetics, and in some cases with product inhibition. For the majority of the analyses, the quantity of the available data were sufficient and the model was constructed with Michaelis-Menten equations. There were however instances where the available data was not enough to fully parameterise parts of a system. In these cases mass action kinetics were used to describe those parts of the system. The models used in this study therefore consists either entirely of Michaelis-Menten kinetics or a combination of Michaelis-Menten kinetics and mass action kinetics.

3.1.1 Michaelis-Menten kinetics

Victor Henri was the first to attempt to describe enzyme kinetics with the aid of an equation. He hypothesised that the rate of an enzymatic reaction is proportional to the concentration of the enzyme-substrate complex. He was however not able to prove this experimentally [42, 56]. Michaelis and Menten were able to prove Henri's hypothesis experimentally and in the process the Michaelis-Menten equation was derived, which to this day is still used to describe enzyme kinetics [56, 68]. In 1925 Briggs and Haldane published the derivation of the Michaelis-Menten equation [16]. In its simplest form, the Michaelis-Menten equation is written as:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

where v is the rate of a reaction, V_{max} is the maximal rate, K_m is the dissociation constant (also referred to as the half saturation constant), and $[S]$ represents the substrate concentration [68, 95]. The Michaelis-Menten equation can be extended and be made reversible to include the effects of product (feedback) inhibition, competitive inhibition, uncompetitive inhibition, noncompetitive inhibition, irreversible inhibition, and allosteric regulation [95].

3.1.2 Hill equation

The Hill equation, derived by Archibald Hill in 1910, is a variation of the Michaelis-Menten equation that describes the effects of allosteric regulation and cooperativity on enzyme kinetics:

$$v = \frac{V_{max}[L]^n}{K^n + [L]^n}$$

where n is the Hill coefficient which is number of ligand binding sites per molecule of enzyme, K^n is the constant including the effects of the K_m and the interaction factors of the allosteric binding sites, and $[L]$ is the ligand concentration [37, 95].

3.1.3 First order kinetics or Mass action kinetics

In the case where the concentration of the substrate is much smaller than the dissociation constant ($[S] \ll K_m$), the linear relationship between the enzyme rate and the substrate concentration can be studied as the Michaelis-Menten equation reduces to the following:

$$v = \frac{V_{max}}{K_m}[S] = k[S]$$

where k is a first-order rate constant. The use of this equation is beneficial when little experimental data is available for model parameterisation as the number of parameters

per reaction is reduced to only one unknown [95].

When it comes to fitting the model to data, there are different methods to choose from. Linear and nonlinear regression can be completed with the use of various methods, depending on the requirements of the system [5]. Details of these methods and their applicability will be discussed next.

3.2 Linear regression

Linear regression is used for the analysis of data that describes the relationship between variables in a linear pattern. The most basic example of such a model, where only one independent variable is involved, is a simple linear regression model [17, 73].

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where the dependent variable or response variable, y, is the observed data, x is the independent predictor or regressor variable and β_0 and β_1 are two parameters, also called regression coefficients [17, 73]. The statistical error in the data is represented by ε. It is assumed that these errors are not correlated, have a mean of zero (and unknown variance σ^2). Multiple linear regression models, as shown below, include more regressor variables and parameters [8, 73].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

Models in which the dependent variable is a nonlinear function of the independent variable(s) can be parameterised with linear regression as long as the dependent variable is a linear function of the parameters [17, 73, 92].

3.3 Nonlinear regression

Nonlinear regression models can be written as:

$$y_i = f(\mathbf{x}_i, \boldsymbol{\theta}) + \varepsilon_i, \qquad i = 1, 2, ..., n$$

where y_i is the measured experimental values and ε_i is normally distributed random error with mean zero and variance σ^2 [18, 73]. The function f is called the expectation function with $\mathbf{x}_i$ a vector of independent variables and $\boldsymbol{\theta}$ a $p \times 1$ vector of parameters [73]. At least one of the derivatives of the function f, with respect to the parameters, depends on at least one of the parameters [73].

3.4 Regression methods

Two often used methods for data fitting are the least squares method and the maximum likelihood method [17]. They are intuitive and computationally easy to implement [83].

3.4.1 Least squares estimation method

The least squares method is predominantly used when determining the goodness of fit of linear as well as nonlinear models [17, 18, 36, 92]. The linear and nonlinear least squares methods are used for parameterising models that are either linear or nonlinear in the parameters [36]. The sum of the squared residual values SSR between the measured data and the model fit are minimised. This method returns the parameter value estimates that minimise the sum of the squared residual values [18, 36, 58, 73, 83, 92].

$$SSR(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^{N} \left[y(t_i) - \hat{y}(t_i, \hat{\boldsymbol{\theta}}) \right]^2$$

where $SSR(\hat{\boldsymbol{\theta}})$ is the sum of the squared residual values, N is the number of measured data points $y(t_i)$ and $\hat{y}(t_i, \hat{\boldsymbol{\theta}})$ are the fitted data points [58, 73].

For nonlinear regression models the sum of the squared residual values can also be expressed as:

$$SSR(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^{N} \left[y(t_i) - f(\mathbf{x}_i, \hat{\boldsymbol{\theta}}) \right]^2$$

where the measured data points $\hat{y}(t_i, \hat{\boldsymbol{\theta}})$ can be expressed as a linear or nonlinear function $f(x_i, \hat{\boldsymbol{\theta}})$ of the variables x_i and parameters $\boldsymbol{\theta}$. $\hat{\boldsymbol{\theta}}$ is the set of parameters that is returned as the best fit [73].

3.4.1.1 Weighted least squares fitting and Chi-square

Weighting of experimental data points during the data fitting process is crucial for ensuring accurate model regression [56, 73]. The sum squared residual values between the model fit and the experimental data can be weighted to an average standard deviation value or each of these sum squared residual values can be weighted individually to the variance of the experimental data points [25, 36, 56, 73]. Weighting the sum squared residual value for each of the experimental data points to the variance of that point results in the Chi-squared error criterion:

$$\chi^2 = \sum_{i=1}^{m} \left[\frac{y(t_i) - \hat{y}(t_i, \hat{\boldsymbol{\theta}})}{\sigma_i} \right]^2$$

The χ^2 values have a Chi-square distribution [23]. These values are scalar goodness of fit measurements. The above χ^2 is known as the objective function during regression fitting [36].

3.4.2 Maximum likelihood method

The maximum likelihood can be calculated by minimising the least squares estimates if the observational noise or error is normally distributed [8, 30, 61]. Both the maximum likelihood and the least squares methods try to determine which set of parameter values will ensure a fitting of the mathematical model closest to the experimental data [20]. Linear and nonlinear least squares methods are similar to the maximum likelihood method when the residual values are normally distributed. Under these conditions minimising the weighted least squares residuals is similar to finding the maximum likelihood estimation of the parameters [67, 126].

The maximum likelihood method is mainly used for statistical models where the mean and the variance are not known and need to be estimated and is thus set as parameters along with the other parameter values. The maximum likelihood method searches over

all possible parameter estimations to find the solution that is most likely to have resulted in the observed data. The set of parameter values that maximises the likelihood of the occurrence of the observed data are returned. Log likelihood can also be used instead of likelihood as the natural logarithm does not affect the maximising of the likelihood [23, 50]. The likelihood function or the log of the likelihood function is differentiated and these partial derivatives are set equal to zero. The Newton-Raphson method or gradient methods can then be used to find the optimal set of parameter values. The Newton-Raphson method might be preferred as it may converge faster than the gradient methods [23]. These methods and other regression algorithms are discussed in more detail below.

3.4.3 Regression algorithms

Least squares methods minimise the sum squared residual value between the model and the experimental data by iteratively changing the parameter values [18, 36, 62]. There are different ways of finding this parameter set. The search methods simply searches for the optimal parameter set over a specified parameter space [18, 25]. The computational time needed can become long when working with more than a few parameters [18]. Gradient methods iterate over the error space and uses the gradient of the error space to find the direction in which the minimum lies. With these methods the optimal parameters are found faster than with the search methods [18, 25]. Gradient methods include Newton's method, Marquardt's method and the Fletcher-Powell method [18]. Most gradient methods are fast, but are hindered by noisy data and can lead to biased parameter estimations and local minima [18].

3.4.3.1 Gauss-Newton

This method starts with an initial estimate for the parameter values $\hat{\theta}_j$ and then uses Taylor series expansions and ordinary least squares estimation to alter the initial parameter estimation. Good initial estimates of the parameter values will ensure that an accurate solution is found faster [8, 36, 92]. A new estimated value $\hat{\theta}_{j+1}$ for the parameter is returned. This process is repeated until the changes in the objective function, equation 3.1, and the estimated parameter values, equation 3.2, are minimal [8, 92].

$$\frac{|SSR(\hat{\theta}_{(j+1)}) - SSR(\hat{\theta}_{(j)})|}{|SSR(\hat{\theta}_{(j)})|} \tag{3.1}$$

$$|\hat{\theta}_{(j+1)} - \hat{\theta}_{(j)}| \tag{3.2}$$

The partial derivative, equation 3.3, is a gradient vector and can be used to determine in which direction the steepest decrease in the sum squared residual values can be found.

$$\frac{\partial SSR(\hat{\theta}_j)}{\partial \theta} \tag{3.3}$$

Overstepping the optimal parameter estimation can be avoided by using the modified Gauss-Newton algorithm. The change in step length i.e. the change from $\hat{\theta}_{(j)}$ to $\hat{\theta}_{(j+1)}$ is modified with a specific value α. There are different methods of choosing α, one of which is step halving, where α is set to $\frac{1}{2}$ [8, 92].

This method is not guaranteed to return the optimal parameter estimations. The success of the Gauss-Newton method depends strongly on the shape of the solution space close to the model solution and how close the initial choice is to the true parameter values [92].

3.4.3.2 Nelder-Mead

This method is categorised as a direct search method. It reduces the size of a simplex to find the coordinates of a minimum in a solution space. A simplex is an N dimensional triangle that represents the estimates of N parameters. With each iteration the largest vertex of the simplex is replaced by a smaller vertex until a minimum solution is found [78].

3.4.3.3 Steepest Descent

During regression fitting this method changes the parameter values in the opposite direction of the objective function's gradient i.e. the parameter values are changed in the downhill direction [36]. This method works well when the objective function is simple. The steepest descent method works fast initially, but then slows down and can take long to converge. This method can be useful when starting parameter optimisation with unknown initial choice [36].

3.4.3.4 Levenberg-Marquardt

This method is used for finding solutions for nonlinear least squares problems [36]. This is a combination of the steepest descent (or gradient descent) and Gauss-Newton methods [5, 36]. The Levenberg-Marquardt method behaves like the Gauss-Newton method when the parameter values are near their optimal solutions, and behaves like the steepest descent method as the parameter values are farther from their optimal solutions [36].

3.4.3.5 Newton-Raphson

Also simply called Newton's method is very similar to the Gauss-Newton method, but with the addition of the use of second partial derivatives 3.4 i.e. the Hessian matrix. This method works only if the elements of the Hessian matrix are positive thus ensuring that the search direction is towards the minimum [18].

$$\frac{\partial^2 SSR(\hat{\theta}_j)}{\partial\theta\partial\theta'} \tag{3.4}$$

3.4.3.6 Quasi-Newton

This method is similar to the Newton-Raphson method that uses the Hessian matrix. For this method the Hessian is replaced by a positive approximation (H_j) of the Hessian. (H_j) is iteratively computed. Second order partial derivatives are no longer used [41].

3.4.3.7 Derivative free methods

These methods include the Secant method, which is similar to the Gauss-Newton method apart from the derivative that are numerically calculated from past iterations [101].

3.4.3.8 Possible issues

Failing to converge might suggest that there is little variance in the sum of the squared residual values near the minimum i.e. the solution space is flat. Starting the fitting process with more accurate initial estimates of the parameter values might solve this problem [92]. An overly complex model with simple data can also fail to converge. Converging to a local minima might be solved by using various initial values or by

including parameter constraints.

Levenberg-Marquardt is a local search method that might not return the global minimum as the optimal set of parameter solutions. Global search methods converge slowly once near a global minimum. A good strategy is to use a global search method to find an initial estimation for the parameter values whereafter a local search method is used to find the optimal global minimum values for the parameters [5].

3.5 Identifiability analysis

Model parameters are considered identifiable if given a model describing the kinetics of a system and a set of experimental data, the parameter values can be uniquely determined [55]. If parameters are identifiable it is possible to determine the parameter values of a model with the use of a specified set of initial concentrations and operating conditions while the fitted model accurately describes a specific set of experimental data [127].

Once the model has been fitted to experimental data, weighted to the variance, a set of parameter estimations are returned. Now one needs to test the identifiability of these parameter estimations. The goal of model fitting to data is not only to estimate parameter values, but to determine parameter estimations that are accurate and identifiable. To test the identifiability of parameter values is a popular method of determining parameter value accuracy [59, 64]. There are different methods of verifying parameter identifiability.

3.5.1 Structural and practical non-identifiability

The non-identifiability of parameters can be classified as either practical or structural non-identifiability. Substrate concentrations that do not fully saturate the enzyme during the experimental procedure can lead to practical non-identifiability. The use of an insufficient number of experiments, that describes the enzyme kinetics properly, can also lead to practical non-identifiability. Structural non-identifiability occurs when all the enzyme kinetics cannot be recreated experimentally, for example when it is not possible to obtain data for all the intermediate substrate and product concentrations in a enzyme pathway with multiple branches. Structural non-identifiability is also seen when experimental data is fitted to a model that does not correctly describe the enzyme kinetics [89]. An improvement in the quality or an increase in the quantity of data used during parameter estimation may solve practical no-identifiability, but will not solve structural non-identifiability [64].

3.5.2 Identifiability analysis and experimental design

Due to time and financial constraints it is not always possible to conduct the full number of experiments needed to construct a model and fully parameterise it. The structure of the model sometimes does not allow for all the parameters to be estimated experimentally. Insufficient information about the system, the complexity of the system as well as the lack of observable data can also contribute to the difficulties of parameter identifiability [34, 58, 59, 127].

The model structure, model parameterisation and experimental design can influence parameter estimation and identifiability [30, 59, 127]. Computer simulations of the system dynamics prior to conducting experiments can prevent the occurrence of practical non-identifiability [30, 64]. There is a distinct link between parameter identifiability and experimental design. An increase in the number of experiments and the variety of experimental conditions will likely increase the number of identifiable parameter estimations [64, 127].

Identifiable parameters are constrained within upper and lower bound confidence intervals, therefore identifiability analysis and the calculation of confidence intervals are closely linked.

3.6 Confidence intervals

When analysing experimental data with a model, much importance lies in calculating confidence intervals for the optimised parameter values [3, 4]. The width of confidence intervals are an indication of the goodness of fit of the model to the data. Wider confidence intervals are an indication of poor parameter estimations where narrow confidence intervals are an indication of a better approximation of the true parameter values [4, 5, 59, 62, 73]. The confidence level (α) used when calculating the confidence intervals represents the probability ($100 * (1 - \alpha)\%$) that an estimated parameter value will fall within the confidence interval. Therefore the choice in confidence level also needs to be considered when evaluating a parameter value's goodness of fit [62, 73]. Sample size can also influence the width of the confidence intervals and larger sample sizes result in narrower confidence intervals [4, 35]. There are two types of confidence intervals: asymptotic confidence intervals and finite sample (also known as likelihood-based) confidence intervals [121].

3.6.1 Asymptotic confidence intervals

Asymptotic confidence intervals can also be calculated with use of the Fisher information matrix [30, 121]. These confidence intervals are most accurate if there is little measurement noise and if the amount of data is large in comparison to the number of parameters [89, 121]. These confidence intervals are accurate when the data is linearly dependent on the parameters [89].

For independent measurement errors that are normally distributed, $\hat{\boldsymbol{\theta}}$ minimises:

$$SSR(\boldsymbol{\theta}) = \sum_{i=1}^{N} \left[\frac{y(t_i, \boldsymbol{\theta}) - \hat{y}(t_i, \hat{\boldsymbol{\theta}})}{\sigma} \right]^2 = Y^T(\boldsymbol{\theta}) Y(\boldsymbol{\theta})$$

where $Y(\boldsymbol{\theta})$ is the vector of residual values between the experimental data and the fitted model values [5, 59]. With this vector of residual values the Jacobian matrix can be calculated which is used to determine the asymptotic confidence intervals.

3.6.1.1 Jacobian matrix

$$\boldsymbol{J}(\boldsymbol{\theta}) = \frac{\partial Y(\boldsymbol{\theta})}{\partial(\boldsymbol{\theta})}$$

$\boldsymbol{J}$ is the $N \times p$ matrix of partial derivatives, called the Jacobian matrix, with N number of observations or measurements and p number of parameters [5, 59, 92]. The Jacobian matrix consists of derivatives of the residual vector with respect to the parameters [29, 36, 53]. The Jacobian matrix can also be seen as the sensitivity matrix [5, 53]. The Jacobian matrix is an indication of the local sensitivity of the fitting function to changes in the parameters [36, 53]. The parameters are only normally distributed if the errors in the data are normally distributed and if the model is a linear function in the parameters and variables [2, 73]. If these conditions are satisfied, the confidence intervals can be calculated with the use of Student's t statistic [2].

Analytical methods of calculating the Jacobian matrix render an accurate result. This method of calculating the Jacobian is not always feasible for large datasets or models of nonlinear form as the sensitivity of the model is considered at each individual datapoint. This method works well for small models, but can be computationally complex and time consuming for larger models. Numerical methods of calculating the Jacobian matrix is therefore better when working with larger models as they are computationally easier to

calculate [5, 29].

3.6.1.2 Divided differences

A more feasible approach for calculating the Jacobian when working with large scale problems is the use of divided differences, also called finite differences or forward differences [5, 29, 36, 49]. Considering the sensitivity of the model fit in terms of datasets and not individual datapoints can be computationally more cost effective for large problems [5].

$$\boldsymbol{J}(\theta)_{ij} = \frac{\partial Y(\theta)}{\partial \theta_j} = \frac{Y(\theta + \partial\theta_j) - Y(\theta)}{\partial \theta_j}$$

where $\partial\theta_j$ is a small perturbation in the j^{th} parameter [5, 29, 36].

The perturbation size can influence the accuracy of the this method [29, 49, 54]. A relatively small perturbation size is needed for the calculation of the derivative values [49]. A perturbation size of between 10^{-3} and 10^{-7} return good estimations of the derivative values [29, 54] with the optimal perturbation size being between 10^{-4} and 10^{-6} [54, 66]. The symbolic manipulation tools of programs such as Mathematica can also be used to calculate the Jacobian matrix, but due to the computational resources needed when working with large models, this approach is better when working with small models [49].

3.6.1.3 Error variance

Along with the Jacobian matrix, the error variance is also needed in the process of calculating asymptotic confidence intervals. During least squares fitting, residual values are calculated and minimised. It is assumed that these residual values (r_i) are normally and independently distributed with a mean (μ) of zero [62, 73], also presented as:

$$\mu(r_i) = \frac{1}{N - m} \sum_{i=1}^{N} r_i$$

The variance (σ^2) in the residual values can be calculated as:

$$\sigma^2(r_i) = \frac{1}{N-m}\sum_{i=1}^{N}(r_i - \mu)^2$$

for N number of data measurements and m number of variables [73, 92].

The mean squared measurement error is calculated after the optimal set of parameter values have been determined [36].

$$\hat{\sigma}^2 = \frac{SSR}{N-p}$$

where p is the number of parameters. The variance and Jacobian matrix is needed to calculate the covariance matrix, which is then needed for calculating asymptotic confidence intervals [73].

3.6.1.4 Covariance matrix

The covariance matrix of the parameter estimations is shown below [62, 73, 92]:

$$\boldsymbol{Cov} = \hat{\sigma}^2(\boldsymbol{J}^T\boldsymbol{J})^{-1}$$

$$\boldsymbol{Cov} = \hat{\sigma}^2(\boldsymbol{H})^{-1}$$

where $\boldsymbol{H}$ is the Hessian matrix, a matrix of second partial derivatives [5, 18].

The diagonal elements of $\boldsymbol{Cov}$ is the variance in the parameters [2, 62, 73] and the other elements are the covariance between the parameters [73]. The square root of Cov_{jj} is the standard error of the parameter, θ_j, if the parameter estimation adheres to Gaussian (normal) distribution [2]. The inverse of Cov_{jj} can therefore be used as the weighting matrix during regression fitting [36]. The weighting matrix is a matrix of diagonal elements where W_{ii} is $\frac{1}{w_1}$. The diagonal elements can also be described as $\frac{1}{\sigma_1^2}$ [36, 73, 117].

Large off-diagonal numbers of the Covariance matrix are an indication of parameter correlations, however this is only an indication of the linear dependancies between two parameters and not between three or more parameters [59].

Cov_{jj} is not always representative of the uncertainty in the parameter θ_j. Cov_{jj} is the true standard error estimate for θ_j if the other parameter values have been fixed to their respective optimal values. Cov_{ij} is the covariance of θ_i and θ_j [2].

If parameters are not correlated they are orthogonal. It is often difficult to determine accurate estimations for parameters, whether they are orthogonal or correlated. Large covariance values (off diagonal elements) (C_{ij}) may be an indication of parameter correlation. The larger the covariance values the poorer the parameter identifiability. This is however not a perfect representation of parameter identifiability as the covariance values only represent the correlation between two parameters and not between one parameter and the rest [59, 73].

3.6.1.5 Correlation coefficients

Correlation coefficients are an indication of the degree to which two parameters are correlated. The correlation coefficients vary between -1.0 and 1.0 where -1.0 is an indication of a 100% negative correlation between the two parameters while 1.0 indicates positive correlation. A correlation coefficient of 0 indicates that there is no correlation between the parameters [75]. The correlation coefficient between two parameters can be calculated by dividing the covariance between the parameters by the product of the standard deviations of the two parameters [5].

$$\rho_{xy} = \frac{Cov(r_x, r_y)}{\hat{\sigma}_x \hat{\sigma}_y}$$

3.6.1.6 Calculating Asymptotic Confidence Intervals

The method described below works only if the residual values adhere to Gaussian distribution. This means that if the squared difference values between the measured data and the model fit are not normally distributed, this method will not return accurate confidence intervals [92]. This method of calculating the confidence intervals does not consider the covariance between pairs of parameters [3]. The asymptotic confidence intervals for the parameters can be calculated as:

$$\theta = \hat{\theta} \pm \delta\sqrt{Cov_{jj}}$$

where the constant, δ, ensures that the confidence intervals are correctly calculated for a specific probability. There are different methods with which to calculate this constant, one of which is the simple Bonferroni method [62, 73], where δ is calculated as follows:

$$\delta = t_{\frac{\alpha}{2r},n-p}$$

where r is the number of independent variables, α is the significance level, n is the number of observations and p is the number of parameters.

The parameters will have joint confidence intervals with a coefficient of at least $1-\alpha$ if confidence intervals of $100(1-\frac{\alpha}{2})$ percent are constructed for each of the parameters.

With the Scheffé S-method δ is calculated as:

$$\delta = \sqrt{rF_{\alpha,r,n-p}}$$

With the maximum modulus t procedure δ is calculated as:

$$\delta = u_{\alpha,r,n-p}$$

where $u_{\alpha,r,n-p}$ is the upper α part of the distribution of the maximum absolute value of r variables. These variables are independent student-t random variables [73].

The maximum modulus t confidence intervals are generally smaller than the Bonferroni confidence intervals and the Scheffé confidence intervals are generally larger than both the Bonferroni and maximum modulus t confidence intervals. The width of the Scheffé confidence intervals are not influenced by the number of independent variables, while the Bonferroni and maximum modulus t confidence intervals increase as more

independent variables are involved. When r is very large the Scheffé confidence intervals will be smaller than the Bonferroni and maximum modulus t confidence intervals [73].

The confidence intervals of parameter estimations can also be calculated by multiplying the standard deviations in the parameters by the Student's t statistic i.e. setting δ equal to the Student's t statistic [3]. The Student's t statistic ($t_{\frac{\alpha}{2},n-p}$) is read from a table or calculated by the programming software used. The significance level (α) will be 0.05 for a 95% confidence level. Half of α and the degrees of freedom are the two values used to read the t statistic from the table. The degrees of freedom are calculated by subtracting the number of parameters (p) from the number of observations (n). The scalar value of the error variance is used and is calculated by dividing the sum of the squared residual values by the degrees of freedom [62, 73, 92].

This is only one method with which confidence intervals can be calculated. Another method is the likelihood-based method or finite sample method.

3.6.2 Likelihood-based confidence intervals

Likelihood based confidence intervals are also called finite sample confidence intervals [89]. As measurement error is a constant factor to consider, likelihood-based confidence intervals are more accurate descriptions of the confidence intervals. This method is more reliable under these circumstances as well as when data with minimal information is used [121]. Advances in computational technology has made it easier to apply likelihood functions and likelihood ratios to a wide range of problems. There are advantages to using likelihood-based methods over using asymptotic methods for calculating confidence intervals [67]. Asymptotic methods depend on symmetric parameter distribution and accurate, useful standard error estimations [109]. Often asymmetric confidence intervals (such as likelihood-based confidence intervals) give a more accurate description of the optimal parameter value estimation and its corresponding confidence interval. It also works well for smaller data sample sizes as the profile likelihood does not require a normal distribution of the parameter estimation [20]. The likelihood-based approach is more reliable in indicating structural non-identifiability than the asymptotic approach [34].

To calculate the confidence regions for multiple parameters the Chi-squared test can be used [92]. A likelihood profile for each parameter is created by fixing each parameter to a value and refitting for all the other parameters. As the possible value for each parameter is varied, the other parameters are fitted for and the likelihood that the data could have been generated by the model is determined. The confidence

interval is thus a range of values of which the likelihood that they could have been generated by the model is close to the likelihood value of the optimal parameter value [20].

3.6.2.1 Log-likelihood and Profile likelihood

The log-likelihood function is minimised in terms of all parameters except for one which is kept fixed at a specific value [20, 59]. For each of the values explored within the possible solution space a maximum log-likelihood value is calculated. This function is called the profile log-likelihood of the parameter values explored. The antilog thereof is the profile likelihood. The profile likelihood method works well when it is challenging to calculate the confidence intervals with other standard methods, for instance when the function is not normally distributed. This method is also not influenced by the size of the data set. Other methods, such as the asymptotic Wald method depends on larger data sets that have normal distributed maximum likelihood estimates [20].

Profile likelihood is used for the identifiability analysis of partially observable models [64]. The model is refitted to the data and a likelihood-based confidence region is created for each parameter as the parameter space for each parameter is explored [20, 64]. The parameter space is explored by changing the parameter values slightly and then refitting for the other parameters [20, 88]. It is also possible to distinguish between structural and practical non-identifiability with the use of the profile likelihood-based approach [61, 64, 89]. This profile likelihood that is generated for each parameter can be subjected to the likelihood ratio test to determine whether the null hypothesis or alternative hypothesis is accepted. The use of the likelihood ratio test is discussed next.

3.6.2.2 Likelihood ratio test

The likelihood ratio of two hypotheses is a measurement of their plausibility, calculated as:

$$\Lambda = \frac{max[L_0]}{max[L_1]}$$

where L_0 is the likelihood under the null hypothesis and L_1 is the likelihood under the alternative hypothesis. The null hypothesis is accepted when $\Lambda > 1$ and the alternative hypothesis is accepted when $\Lambda < 1$ [50]. The likelihood ratio test works on the principal of estimating a parameter with and without a certain constraint and then comparing the estimation criteria [51]. The log likelihood test statistic can be calculated as:

$$S_L(\theta) = n\Big(logL_0 - logL_1\Big)$$

where the number of observations (data points) are represented by n [51]. In cases when the null hypothesis is rejected (i.e $\Lambda < 1$), the null hypothesis is a special case of the alternative hypothesis. Then the likelihood ratio test statistic follows asymptotic Chi-squared distribution with q degrees of freedom where q is the difference in the number of free parameters between the the null and alternative hypotheses. As the number of observations approach infinity, the likelihood ratio and the log-likelihood ratio follow a chi-square distribution with 1 degree of freedom. The significance level is inferred by comparing the likelihood test statistic value to Chi-squared distribution with 1 degree of freedom. The Chi-squared distribution is used to test the significance of the likelihood ratio test statistic [50, 51, 67, 109].

3.6.2.3 Calculating the likelihood-based confidence intervals

Calculating confidence intervals for a parameter based on the log likelihood ratios are done by calculating all the values of the parameter for which the log likelihood $logL_1$ is within a given variation from the optimal parameter value likelihood $logL_0$. $\hat{\boldsymbol{\theta}}$ maximises the profile likelihood L_0. Confidence interval $100(1-\alpha)\%$ for the parameter $\boldsymbol{\theta}$ consists of all the possible values of the parameter for which the null hypothesis is not rejected [89].

The confidence interval for parameter $\boldsymbol{\theta}$ is the set of all values of $\boldsymbol{\theta}$ such that the following is true:

$$S_L(\boldsymbol{\theta}) < \chi^2_{1-\alpha,df}$$

where $\chi^2_{1-\alpha,df}$ is the $1-\alpha$ quantile of the Chi-squared distribution with df degrees of freedom [61, 67, 120]. The upper and/or lower confidence interval limit values are the values of θ for which the following is true:

$$n\Big(logL_0 - logL_1\Big) = \chi^2_{1-\alpha,df} \tag{3.5}$$

The profile likelihood of each parameter and their respective likelihood-based confidence intervals can be presented as profile likelihood plots. The graphical representation of these results are often easier to interpret.

3.6.3 Profile likelihood plots

The likelihood-based confidence intervals are often better presented in the form of profile likelihood plots. These plots are created by fixing one of the parameter values and refitting for the other parameters. The parameter is fixed to different values across the solution space. A sum squared residual value is returned for each fit. A smaller SSR indicates a better parameter estimation, while a larger SSR value is indicative of a worse fit. If the SSR values increase with a change in the fixed parameter values, the other parameters are not able to compensate for the change in the fixed parameter value.

Plotting each value in the parameter solution space against its log likelihood estimate gives a graphical representation of the confidence intervals. The threshold in the profile likelihood ($\chi^2_{1-\alpha,df}$) can also be indicated, making it easy to distinguish whether parameters have confidence intervals that have upper or lower limits.

Practical non-identifiability is indicated by only either a upper or lower constrained confidence interval and structural non-identifiability is indicated by confidence intervals that are neither upper or lower constrained. It is therefor easy to identify practical or structural non-identifiability by considering the shape of the profile likelihood plots [61, 89, 109].

Likelihood-based confidence intervals can be calculated with the use of sparse data and do not require symmetric parameter distribution. This method of calculating confidence intervals can in some cases be computationally expensive. In contrast, calculating asymptotic confidence intervals can require less computational power, however, this method requires normal distribution of parameter estimations and large sample sizes of data.

Throughout this study the likelihood-based method of calculating confidence intervals have been used. Profile likelihood plots have been created for all parameters to indicate identifiability and the likelihood ratio test have been used to calculate confidence intervals.

3.6.3.1 Theoretical example

The example shown in this section illustrates how profile likelihood plots can be used to aid experimental design as well as to distinguish between structural and practical

non-identifiability (as addressed in section 3.5.1). Consider the case where one wishes to parameterise the dynamics of a certain enzyme of which the kinetics can be described by the irreverisible Michaelis-Menten equation:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

If another equation, which differs vastly from the Michaelis-Menten equation, shown above, is used to determine the parameter values, the parameter values will most likely be structurally non-identifiable; irrespective of the quantity and quality of the experimental data used for model fitting. Creating a profile likelihood plot of either of the parameters V_{max} and K_m will result in a plot similar to the example Figure 3.1.

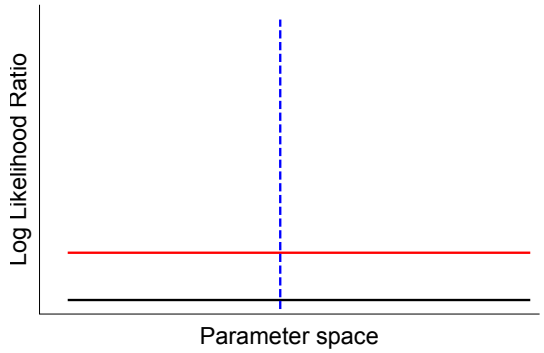


Figure 3.1: **Theoretical example:** Profile likelihood plot of a structurally non-identifiable parameter value. The black line indicates the profile likelihood, the red line indicates the 95% confidence threshold, and the dashed blue line indicates the parameter value estimation.

Using the correct equation for model parameterisation, but the experimental data used is lacking in either quantity or quality could result in parameter values being practically non-identifiable. The example dataset shown in Figure 3.2A is an example of data lacking in quantity. The initial rates of these two time courses are below the K_m value, as can be seen in Figure 3.2B. The dynamics of the system can very well be approximated with linear kinetics and this may result in a profile likelihood shape similar to the example Figure 3.3.

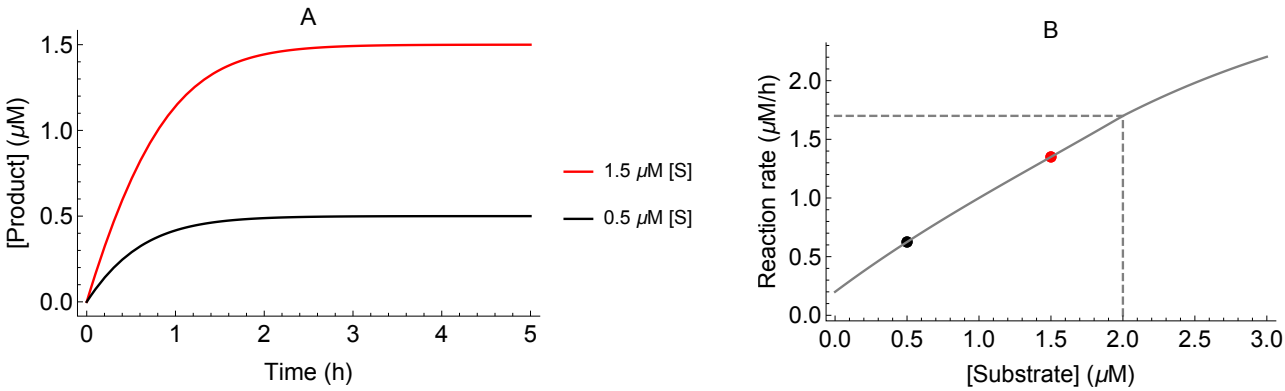


Figure 3.2: **Theoretical example:** (A) data lacking in quantity resulting in practical non-identifiability. (B) The initial rates of the time course data is below the K_m value and thus linear kinetics can be used to approximate the parameter values.

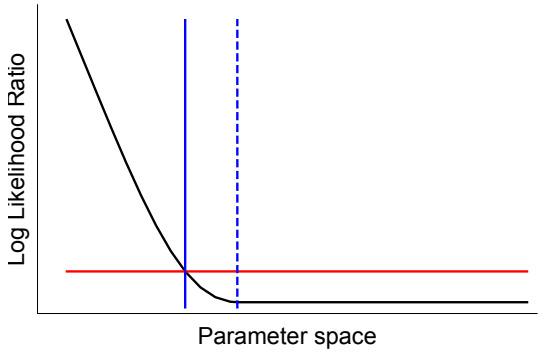


Figure 3.3: **Theoretical example:** Profile likelihood plot of a practically non-identifiable parameter value. The black line indicates the profile likelihood, the red line indicates the 95% confidence threshold, the dashed blue line indicates the parameter value estimation, and the solid blue line indicates the one boundary of the 95% confidence interval.

Identifiable parameter value estimations can be determined with the use of the correct model, describing the system dynamics, as well as the use of sufficient data. If, for example, time course datasets that saturate the enzyme is added to the two datasets shown above, the identifiability of the parameters could be improved. Adding datasets of which the initial rates are above the K_m value (see Figures 3.4A and 3.4B) could result in identifiable parameter values, with profile likelihood plots similar to the example Figure 3.5.

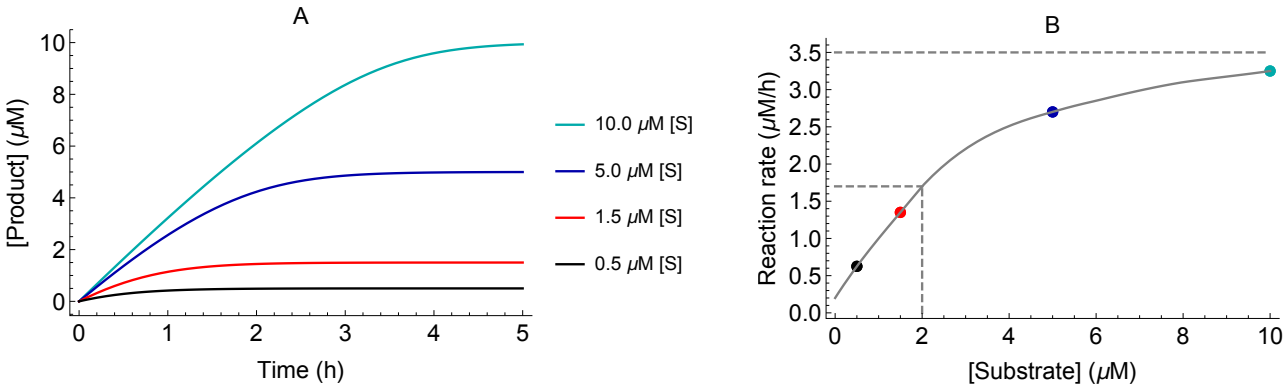


Figure 3.4: **Theoretical example:** (A) data resulting in identifiable parameter value estimations. (B) The initial rates of the time course data follows the shape of the full Michaelis-Menten curve.

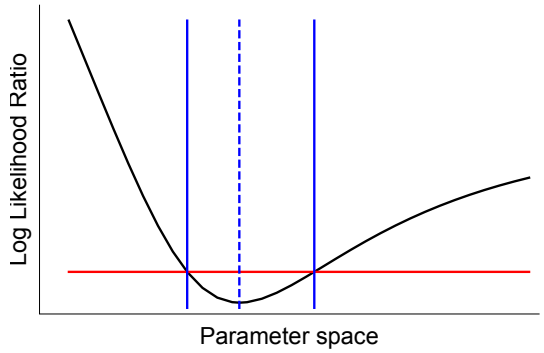


Figure 3.5: **Theoretical example:** Profile likelihood plot of an identifiable parameter value. The black line indicates the profile likelihood, the red line indicates the 95% confidence threshold, the dashed blue line indicates the parameter value estimation, and the solid blue lines indicate the boundaries of the 95% confidence interval.

The following section contains an overview of metabolic control analysis techniques used in the subsequent research chapters.

3.7 Metabolic control analysis

Metabolic Control Analysis (MCA) aids in the understanding of the effects of local changes on systemic properties. In essence MCA can be seen as sensitivity analysis of a system. MCA allows one to determine which parts of a biochemical system has the most control over certain properties of the system. The main use of MCA analysis is to quantify the change in the steady state variables and fluxes with variation in parameters or reaction rates [44, 100].

The elasticity, response, and control coefficients need to be determined. MCA is usually for the analysis of a system under steady state conditions, thus all the coefficients need to be defined for a specific steady state.

3.7.1 Control Coefficients

The flux control coefficient is an indication of how strongly the rate of reaction k controls the steady state flux [40, 44, 46, 57]:

$$C_{v_k}^{J_j} = \frac{\partial J_j / \partial p_k}{\partial v_k / \partial p_k}$$

where p_k is a parameter that only affects the rate of reaction k (v_k). This is the non-normalised form of the flux control coefficient. $\partial v_k / \partial p_k$ is evaluated with the parameter values and concentrations associated with the specific steady state.

The change in the flux and the rate is usually normalised by dividing it by the respective reference values at steady state:

$$C_{v_k}^{J_j} = \frac{v_k}{J_j} \frac{\partial J_j / \partial p_k}{\partial v_k / \partial p_k}$$

The concentration control coefficient is an indication of how strongly the rate of reaction (k) controls the metabolite concentration (s_j):

$$C_{v_k}^{s_j} = \frac{\partial s_j / \partial p_k}{\partial v_k / \partial p_k}$$

3.7.2 Elasticity Coefficients

Elasticity coefficients give a quantitative measurement of the effect of metabolites or parameters on the rate of an enzyme [100]. These effects can be due to a change in the concentration of the enzyme or the metabolites, or a change in a parameter value. These coefficients are scaled partial derivatives of the kinetic rate equations with respect to

metabolite concentrations or parameter values [46, 100].

$$\mathcal{E}_p^{v_k} = \frac{\partial v_k / v_k}{\partial p / p}$$

3.7.3 Response Coefficients

Response coefficients quantify the change in the flux (J) relative to a perturbation in a parameter value (p) [46, 57]:

$$R_p^J = \frac{\partial J / J}{\partial p / p}$$

The relationship between control, elasticity, and response coefficients is given by [46]:

$$R_p^J = C_k^J \cdot \mathcal{E}_p^{v_k}$$

3.7.4 Combined Response Theorem

Concentration response coefficients describe the steady state response in a metabolite concentration (s_j) to variation in the value of a parameter (p_k) [44].

$$R_{p_k}^{s_j} = \frac{\partial s_j}{\partial p_k}$$

3.7.5 Summation and Connectivity Theorems

These theorems give insight into the link between the global properties of a pathway and the local properties of individual enzymes. Therefore it is possible to determine the control distribution over the enzymatic pathway which, in turn, relates the system behaviour to specific branches of the pathway [100].

The summation theorem shows quantitively how the control over the flux is distributed in the system [46, 57]:

$$\sum_{k=1}^{n} C_k^J = C_1^J + C_2^J + \cdots + C_n^J = 1$$

The summation theorem also applies to the steady state concentrations of the metabolites, however, the concentration control coefficients will sum to 0 and not 1 like the flux control coefficients [40]:

$$\sum_{k=1}^{n} C_k^{s_j} = C_1^{s_j} + C_2^{s_j} + \cdots + C_n^{s_j} = 0$$

The flux connectivity theorem shows the relationship between the control coefficients and the elasticity coefficients for the reactions of a system affected by a metabolite (s_j) [57]:

$$\sum_{k=1}^{n} C_k^J \mathcal{E}_{s_j}^{v_k} = C_1^J \mathcal{E}_{s_j}^{v_1} + C_2^J \mathcal{E}_{s_j}^{v_2} + \cdots + C_n^J \mathcal{E}_{s_j}^{v_n} = 0$$

The control coefficients can also be calculated from the elasticity coefficients [44]:

$$C = \mathcal{E}^{-1}$$

These MCA techniques have been used throughout all three research chapters in this study. The following section describes the computational application of the methods discussed in this chapter and the development of the `IdentifiabilityAnalysis` Mathematica package.

3.8 Computational application: Mathematica package

This section is a description of the development of the Mathematica add-on package, `IdentifiabilityAnalysis`. This package has been developed to simplify the process of

linear or nonlinear regression fitting of an ordinary differential equation (ODE) based model to experimental data. The model fitting technique used in this study, which has therefore been included in this packages, is nonlinear least squares regression with the Nelder-Mead regression method. The use of the functions of this package is not limited to nonlinear least squares regression problems; although all of the models developed in this study were nonlinear, the functions in the package are compatible with linear problems.

The package also includes testing the identifiability of the newly determined parameter values with the likelihood ratio test and creating profile likelihood plots which give a visual depiction of the parameter identifiability and confidence intervals. Although these methods of model parameterisation and identifiability analysis are established and widely used in various disciplines, Mathematica functions that complete these analyses, especially the likelihood ratio test and creating the profile likelihood plots, have, to the best of our knowledge, not been made available for public use. This package allows the user to parameterise a model, complete identifiability analysis and create profile likelihood plots with minimal effort and little additional coding needed. The package is available on request and will be made available on GitHub once the necessary accompanying help documentation has been developed. The package includes five functions:

`objectiveFunction[modelAndData_, parListNames_, parListVals_]`
This function finds the sum squared residual value between the model and the data, `modelAndData`, weighted to the variance with the parameters, `parListNames`, set to the values, `parListVals`. The format of `modelAndData` should be a nested list containing a list of ODEs, a list of initial values of the variables, a list of the variables, and the data including the data variance. This nested list can include one entry if only one dataset is used or multiple entries for the case where the model should be fitted to a number of experimental datasets. A single value is returned by this Mathematica function. The objective function in this method is the Chi-squared error criterion discussed in section 3.4.1.1.

`minimizeObjectiveFunction[modelAndData_, parList_, minima_, maxima_]`
This function finds the solution of the parameters, `parList`, that returns the smallest sum squared residual value between the model and the data, `modelAndData`, weighted to the variance. The parameters, `parList`, are constrained by the lower limits, `minima`, and the upper limits, `maxima`. This function uses the built-in Mathematica functions `NDSolve` and `NMinimize` to minimise the objective function. The Nelder-Mead regression method in section 3.4.3.2 is used by default.

`profileLikelihood[modelAndData_, parList_, minima_, maxima_, lowlim_, uplim_, counter_]`
This method returns the profile likelihood values for the parameters, `parList`, of the model with the accompanying data, `modelAndData`. One parameter value is varied from the value `lowlim` times smaller than the parameter value to the value `uplim` times larger than the parameter value for `counter` number of intervals, while a solution for the other parameters that returns the smallest sum squared residual value between the model and the data, weighted to the variance. The parameters, `parList`, are constrained by the lower limits, `minima`, and the upper limits, `maxima`. This method tests for the identifiability of the parameter values determined by the function `minimizeObjectiveFunction`. For each parameter a list of values is returned representing the profile likelihood of that parameter, as discussed in section 3.6.2.1.

`likelihoodRatioTest[dataAll_, pls_, thresholdVal_Real]`
This function returns the values of the likelihood ratio test of the profile likelihood values, `pls`, for the data, `dataAll`, used during fitting, for the Chi-square distribution value, `thresholdVal`. The user inputs the Chi-square distribution value that corresponds to one degree of freedom and the desired confidence level. The log likelihood test statistic, mentioned in section 3.6.2.1, is then calculated according to the number of data points, taken from `dataAll`, and the profile likelihood values, `pls`.

`likelihoodRatioPlot[dataAll_, pls_, thresholdVal_Real]`
This function returns the plot of the likelihood ratio test of the profile likelihood values, `pls`, for the data, `dataAll`, used during fitting, for the Chi-square distribution value, `thresholdVal`. The theory of these profile likelihood plots are discussed in section 3.6.3.

The `IdentifiabilityAnalysis` package was applied to complete all subsequent analyses in this study.

Chapter 4

A comparative analysis of 17-OHPROG and A4 synthesis in ovine and Angora goat models

South Africa is the largest exporter of mohair, currently supplying about 55% of the total worldwide mohair stock [74]. The industry suffers the loss of young, newly shorn Angora goats in cold weather. This phenomenon has been studied extensively and it was found that the Angora goats suffer from altered adrenal function, resulting in hypocortisolism. Sufficient cortisol synthesis would result in higher blood glucose levels which in turn allows the animals to survive the cold spells. It was originally hypothesised that reduced adrenal function in the Angora goats leads to the hypocortisolism seen in these animals [118]. Later it was found that the reduced cortisol synthesis may be due to increased CYP17 function relative to 3βHSD function rather than overall reduced adrenal function [28]. The CYP17 enzyme catalyses three reactions in the Δ^4 pathway and two reactions in the Δ^5 pathway. Considering the interlinked dynamics of the Δ^4 and Δ^5 pathways, it is difficult to predict the effect of this altered enzymatic function on overall steroid synthesis without a computational model of Angora steroidogenesis which has to date not been developed. Such a model would aid in gaining a better understanding as to how exactly this increased CYP17 activity leads to the hypocortisolism seen in Angora goats.

The first aim of this chapter is to establish a reference model for healthy cortisol synthesis. Steroidogenesis in the Merino sheep is used for the development of the reference model. The second aim of this chapter is to develop a model that describes the altered adrenal function of the Angora goat. The ovine model is used as a basis for the development of the Angora model. A comparative metabolic control analysis of the two models shows the differences in the steroidogenic function of the two species. We conclude that the increased CYP17 activity in the Angora species causes the Δ^5 pathway

to be favoured with high flux through this pathway even under conditions simulating a cold stress response.

4.1 Parameterisation and identifiability analysis of the ovine Δ^4 and Δ^5 pathways

Both the CYP17 and 3βHSD enzymes contribute to the regulation of steroidogenesis [27, 28, 38, 39]. Here we expand on the work by E. van Schalkwyk and R. Conradie at Stellenbosch University, to establish the ovine model as a control model (unpublished results). In their independent studies, experimental data of metabolite time traces were collected whereafter data fits and some model validation was done. The model validation entailed simulating the outcome of time course data not used during the model construction process. The model is parameterised and validated with experimental data collected in a reconstituted system with the enzymes CYP17 and 3βHSD expressed in transfected COS-1 cells. In this section, the work of Van Schalkwyk and Conradie is combined by pooling the ovine experimental data and reparameterising the model. Thereafter the model is validated in the same manner as mentioned above. Identifiability analyses of all parameters of the Δ^4 and Δ^5 pathways are done, as well as sensitivity analysis. Neither identifiability analysis nor sensitivity analysis have been conducted on this system previously.

First, the experimental data collected by Van Schalkwyk and Conradie were combined and the partial steroidogenic model was reparameterised. All experimental data used were time course data. These datasets span the Δ^4 and Δ^5 branches of the steroidogenic pathway which include the conversion of PREG to PROG, 17-OHPREG, 17-OHPROG, 16-OHPROG, DHEA, and A4 by the enzymes 3βHSD and CYP17, shown in Figure 4.1. The model does not take into account in which zones of the adrenal cortex the steroids are synthesised. The model works with the average values of the steroids in the different zones and assumes that the steroid hormones diffuse out of the cells once they are synthesised.

After our initial reparameterisation of the model, an identifiability analysis of all the parameters for both enzymes was completed. Although the model has previously been parameterised, the identifiability of the parameters were not tested. The identifiability analysis of the preliminary model parameterisations showed non-identifiability of some parameters; the experimental data used were insufficient to obtain full parameter identifiability. The non-identifiable parameters were those describing the kinetics of CYP17. To improve the identifiability, the pool of data needed to be extended and

additional datasets were included in the study (unpublished data collected by R. Louw). These datasets included time courses initiated with higher starting concentrations of the metabolites PREG and PROG to saturate the enzyme, as well as time course data initiated with the intermediate 17-OHPREG, with which only the conversion of 17-OHPREG to DHEA was measured. The model was parameterised again, with the additional data, and the identifiability of the parameters were tested. A summary of the experimental datasets is shown in Tables 4.1 to 4.4. The datasets were grouped into parameterisation datasets (Tables 4.1 and 4.2) and validation datasets (Tables 4.3 and 4.4).

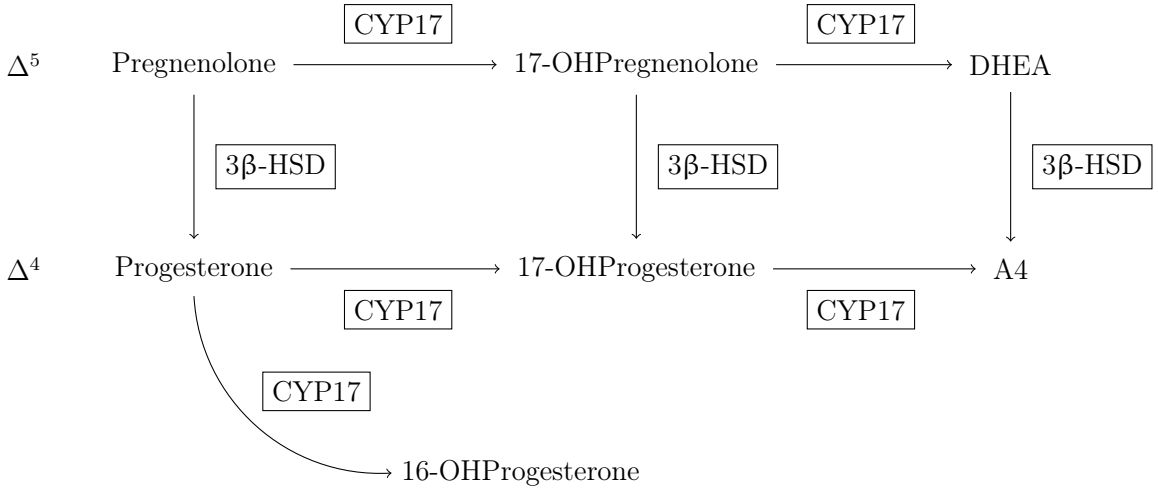


Figure 4.1: **Schematic representation of the Δ^4 and Δ^5 branches of the steroid hormone biosynthesis pathway.**

The majority of the datasets were used for parameterisation, while only a few were used for validation to ensure that the most information is assigned towards finding accurate parameter estimations. The grouping of parameterisation and validation datasets that returned the maximum number of identifiable parameter values were used. The process of cross-validation, as defined in [33], was used to divide the data into either the model construction data or model validation data groups. The validation datasets were not used for model parameterisation, but rather used to test the model. After parameterisation, the model was given the initial conditions under which the validation datasets were collected and asked to predict the outcome of the time course experiments.

Table 4.1: **3βHSD Model parameterisation data**: Initial concentrations of substrates, intermediates and products of the 3βHSD experimental assays completed by Van Schalkwyk. All assays were completed in triplicate [119].

Assay Time	Initial substrate(s)	Metabolites measured	Figure
7h	5 μM PREG, 5 μM PROG	PREG, PROG	Fig. 4.2 A
7h	2 μM PREG	PREG, PROG	Fig. 4.2 B
7h	5 μM PREG	PREG, PROG	Fig. 4.2 C
7h	3 μM PREG, 3 μM PROG	PREG, PROG	Fig. 4.3 G
7h	3 μM PREG	PREG, PROG	Fig. 4.4 B
7h	1.88 μM 17-OHPREG, 2.12 μM 17-OHPROG	17-OHPREG, 17-OHPROG	Fig. 4.2 D
7h	2 μM 17-OHPREG	17-OHPREG, 17-OHPROG	Fig. 4.3 A
7h	3.38 μM 17-OHPREG, 6.62 μM 17-OHPROG	17-OHPREG, 17-OHPROG	Fig. 4.3 B
7h	5 μM 17-OHPREG	17-OHPREG, 17-OHPROG	Fig. 4.3 C
7h	3 μM 17-OHPREG, 3 μM 17-OHPROG	17-OHPREG, 17-OHPROG	Fig. 4.3 H
7h	3 μM 17-OHPREG	17-OHPREG, 17-OHPROG	Fig. 4.4 C
7h	2 μM DHEA	DHEA, A4	Fig. 4.3 D
7h	5 μM DHEA	DHEA, A4	Fig. 4.3 E
7h	5 μM DHEA, 5 μM A4	DHEA, A4	Fig. 4.3 F
7h	3 μM DHEA, 3 μM A4	DHEA, A4	Fig. 4.4 A
7h	3 μM DHEA	DHEA, A4	Fig. 4.4 D

Table 4.2: **CYP17 Model parameterisation data**: Initial concentrations of substrates, intermediates and products of the CYP17 experimental assays completed by Van Schalkwyk and R. Louw. The first four assays were completed by R. Louw (unpublished results) and the last three assays were completed by Van Schalkwyk [119]. All assays were completed in triplicate.

Assay Time	Initial substrate(s)	Metabolites measured	Figure
12h	2.5 μM PREG, 1 μM PROG	PREG, PROG, 17-OHPREG, 17-OHPROG, DHEA, A4	Fig. 4.5 A & Fig. 4.5 B
6h	5 μM PREG	PREG, 17-OHPREG, DHEA	Fig. 4.5 C
12h	5 μM 17-OHPREG	17-OHPREG, DHEA	Fig. 4.5 D
12h	1.5 μM PROG	PROG, 17-OHPROG, 16-OHPROG, A4	Fig. 4.5 E
12h	2 μM PROG	PROG, 17-OHPROG, 16-OHPROG, A4	Fig. 4.5 F
12h	3 μM PROG	PROG, 17-OHPROG, 16-OHPROG, A4	Fig. 4.5 G
6h	5 μM PROG	PROG, 17-OHPROG, A4	Fig. 4.5 H

Table 4.3: **3βHSD Model validation data**: Initial concentrations of substrates, intermediates and products of the 3βHSD experimental assays completed by Van Schalkwyk. All assays were completed in triplicate [119].

Assay Time	Initial substrate(s)	Metabolites measured	Figure
7h	2 μM PREG, 2 μM PROG	PREG, PROG	Fig. 4.6 A
7h	5 μM 17-OHPREG, 5 μM 17-OHPROG	17-OHPREG, 17-OHPROG	Fig. 4.6 C
7h	1.15 μM 17-OHPREG, 2.85 μM 17-OHPROG	17-OHPREG, 17-OHPROG	Fig. 4.6 D
7h	2.23 μM DHEA, 1.77 μM A4	DHEA, A4	Fig. 4.6 B

Table 4.4: **CYP17 Model validation data**: Initial concentrations of substrates, intermediates and products of the CYP17 experimental assays completed by Van Schalkwyk, Conradie, and R. Louw. The first assay was completed by Conradie (unpublished results), the second assay was completed by R. Louw (unpublished results), and the third assay was completed by Van Schalkwyk [119]. All assays were completed in triplicate.

Assay Time	Initial substrate(s)	Metabolites measured	Figure
13h	2 μM PREG	PREG, 17-OHPREG, DHEA	Fig. 4.7 A
12h	4 μM PREG, 1 μM PROG	PREG, PROG, 17-OHPREG, 17-OHPROG, DHEA, A4	Fig. 4.7 B & Fig. 4.7 C
12h	3 μM PREG, 3 μM PROG	PREG, PROG, 17-OHPREG, 17-OHPROG, DHEA, A4	Fig. 4.7 D & Fig. 4.7 E

4.1.1 Model parameterisation and validation

The ovine model is described as rate equations 4.1 and 4.2 and a set of ordinary differential equations 4.3. Michaelis-Menten kinetics are used to describe CYP17 and 3βHSD. All reactions catalysed by these enzymes are irreversible. CYP17 is not affected by product inhibition, but 3βHSD is, this is reflected in the denominators of the rate equations. The regression method used is weighted least squares fitting of the model to the experimental data as stated on page 20. The function `minimizeObjectiveFunction` of the Mathematica add-on package `IdentifiabilityAnalysis` was used. All experiments were completed in triplicate, therefore the standard deviation of each of the experimental data points were used for weighting the individual sum of the squared residual values. The experimental design and experimental conditions under which the data were collected is explained in [119]. Figures 4.2 to 4.4, and Figure 4.5 and shows the data used for the parameterisation of the enzymes 3βHSD and CYP17 respectively. These plots show the experimental data points as well as the fit of the model to the data that corresponds to the fitted solution for the parameter values. Poor fits of the model to some of the experimental data can be seen in these figures. However, with the limited amount of data available for model construction, the choice was made not to exclude any data as outlier datasets. Figures 4.6 and 4.7 show the validation data used to test the newly parameterised model. The model was asked to predict the outcome of these experimental conditions, given the initial metabolite concentrations. The validation of the 3βHSD parameter values show accurate model predictions of the data, however, in Figure 4.7 B an over estimation of the initial reaction is shown. The model predicts an initial influx of substrate as too high. However, considering Figures 4.7 A, C, D, and E, which show accurate model predictions, the model is able to predict the majority of the

validation data accurately.

4.1.1.1 CYP17 rate equations

$$v_1 = \frac{Vmax_1\left(\frac{[PREG]}{Km_{Preg}}\right)}{denominator_{CYP17}}$$

$$v_2 = \frac{Vmax_2\left(\frac{[17OHPREG]}{Km_{17OHPreg}}\right)}{denominator_{CYP17}}$$

$$v_3 = \frac{Vmax_3\left(\frac{[PROG]}{Km_{Prog}}\right)}{denominator_{CYP17}}$$

$$v_4 = \frac{Vmax_4\left(\frac{[PROG]}{Km_{Prog}}\right)}{denominator_{CYP17}}$$

$$v_5 = \frac{Vmax_5\left(\frac{[17OHPROG]}{Km_{17OHProg}}\right)}{denominator_{CYP17}} \tag{4.1}$$

$$denominator_{CYP17} = 1 + \frac{[PREG]}{Km_{Preg}} + \frac{[PROG]}{Km_{Prog}} + \frac{[17OHPREG]}{Km_{17OHPreg}} + \frac{[17OHPROG]}{Km_{17OHProg}}$$

4.1.1.2 3βHSD rate equations

$$v_6 = \frac{Vmax_6\left(\frac{[PREG]}{Km_{Preg}}\right)}{denominator_{3\beta HSD}}$$

$$v_7 = \frac{Vmax_7\left(\frac{[17OHPREG]}{Km_{17OHPreg}}\right)}{denominator_{3\beta HSD}}$$

$$v_8 = \frac{Vmax_8\left(\frac{[DHEA]}{Km_{DHEA}}\right)}{denominator_{3\beta HSD}} \tag{4.2}$$

$$denominator_{3\beta HSD} = 1 + \frac{[PREG]}{Km_{Preg}} + \frac{[PROG]}{Km_{Prog}} + \frac{[17OHPREG]}{Km_{17OHPreg}} + \frac{[17OHPROG]}{Km_{17OHProg}}$$

$$+ \frac{[DHEA]}{Km_{DHEA}} + \frac{[A4]}{Km_{A4}}$$

4.1.1.3 Ordinary Differential Equations

$$\begin{aligned}
\frac{d}{dt}[PREG] &= -v_1 - v_6 \\
\frac{d}{dt}[17OHPREG] &= v_1 - v_2 - v_7 \\
\frac{d}{dt}[DHEA] &= v_2 - v_8 \\
\frac{d}{dt}[PROG] &= v_6 - v_3 - v_4 \\
\frac{d}{dt}[17OHPROG] &= v_3 + v_7 - v_5 \\
\frac{d}{dt}[16OHPROG] &= v_4 \\
\frac{d}{dt}[A4] &= v_5 + v_8
\end{aligned} \tag{4.3}$$

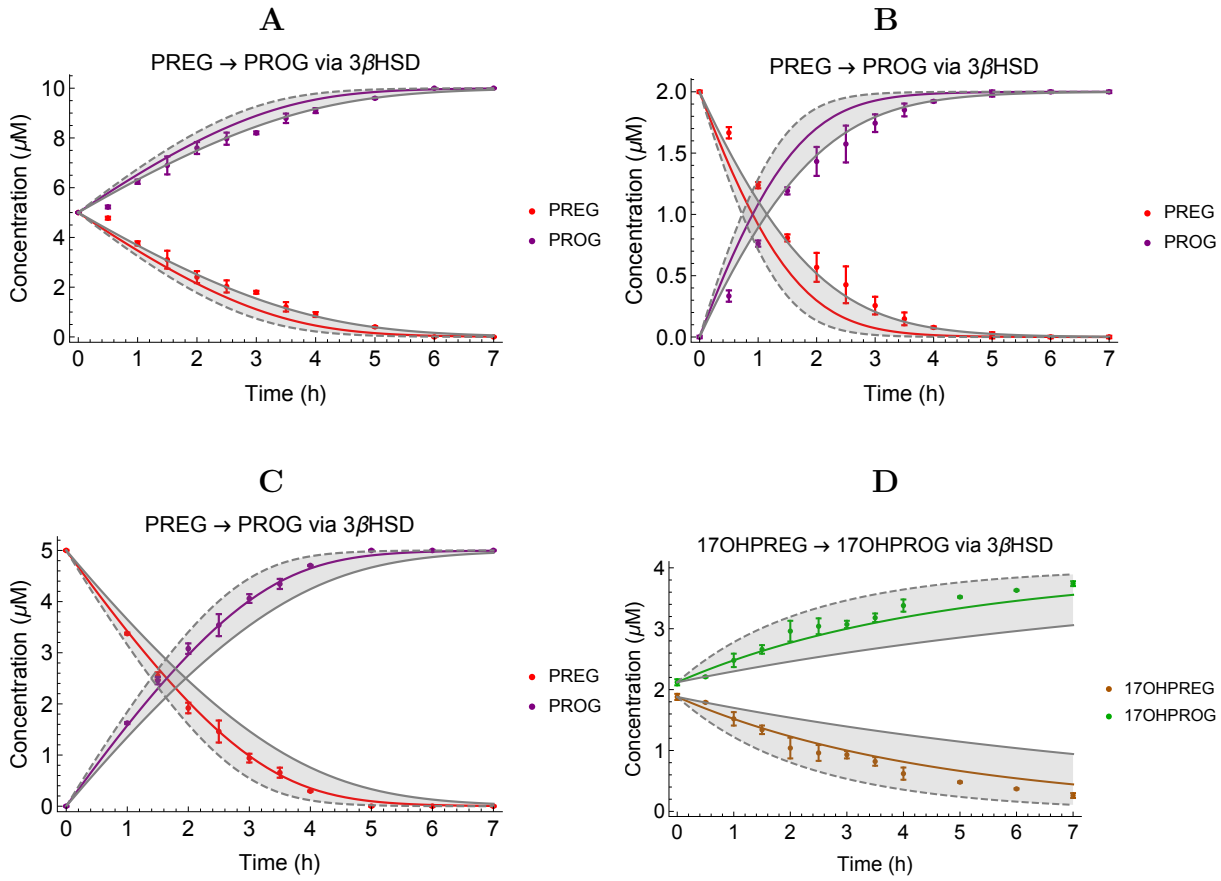


Figure 4.2: **Model fit to 3βHSD data**. Model fit (solid lines) to the experimental data (points) used for the parameterisation of 3βHSD parameters. The 95% confidence intervals are indicated by the grey areas.

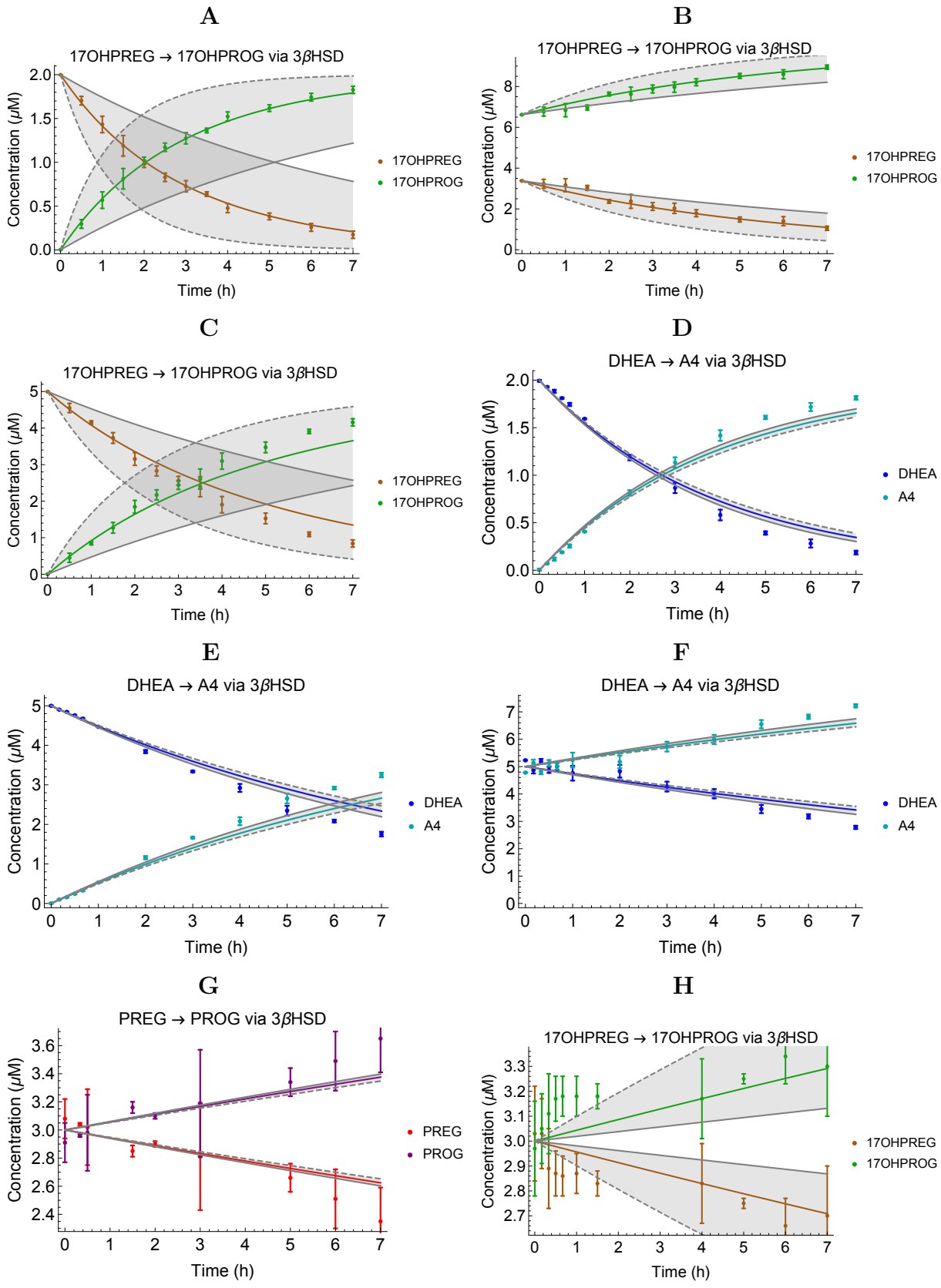


Figure 4.3: **Model fit to 3βHSD data (continued)**. Model fit (solid lines) to the experimental data (points) used for the parameterisation of 3βHSD parameters. The 95% confidence intervals are indicated by the grey areas.

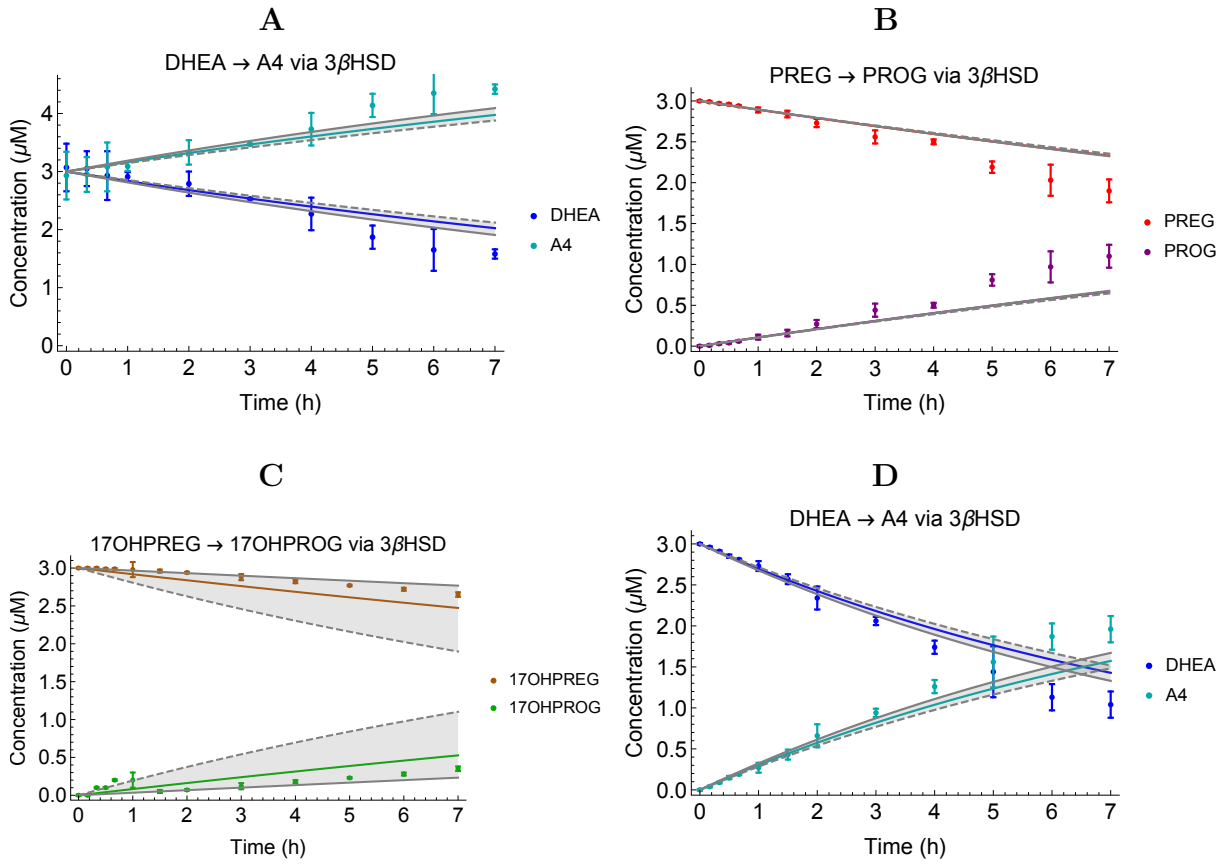


Figure 4.4: **Model fit to 3βHSD data (continued)**. Model fit (solid lines) to the experimental data (points) used for the parameterisation of 3βHSD parameters. The 95% confidence intervals are indicated by the grey areas.

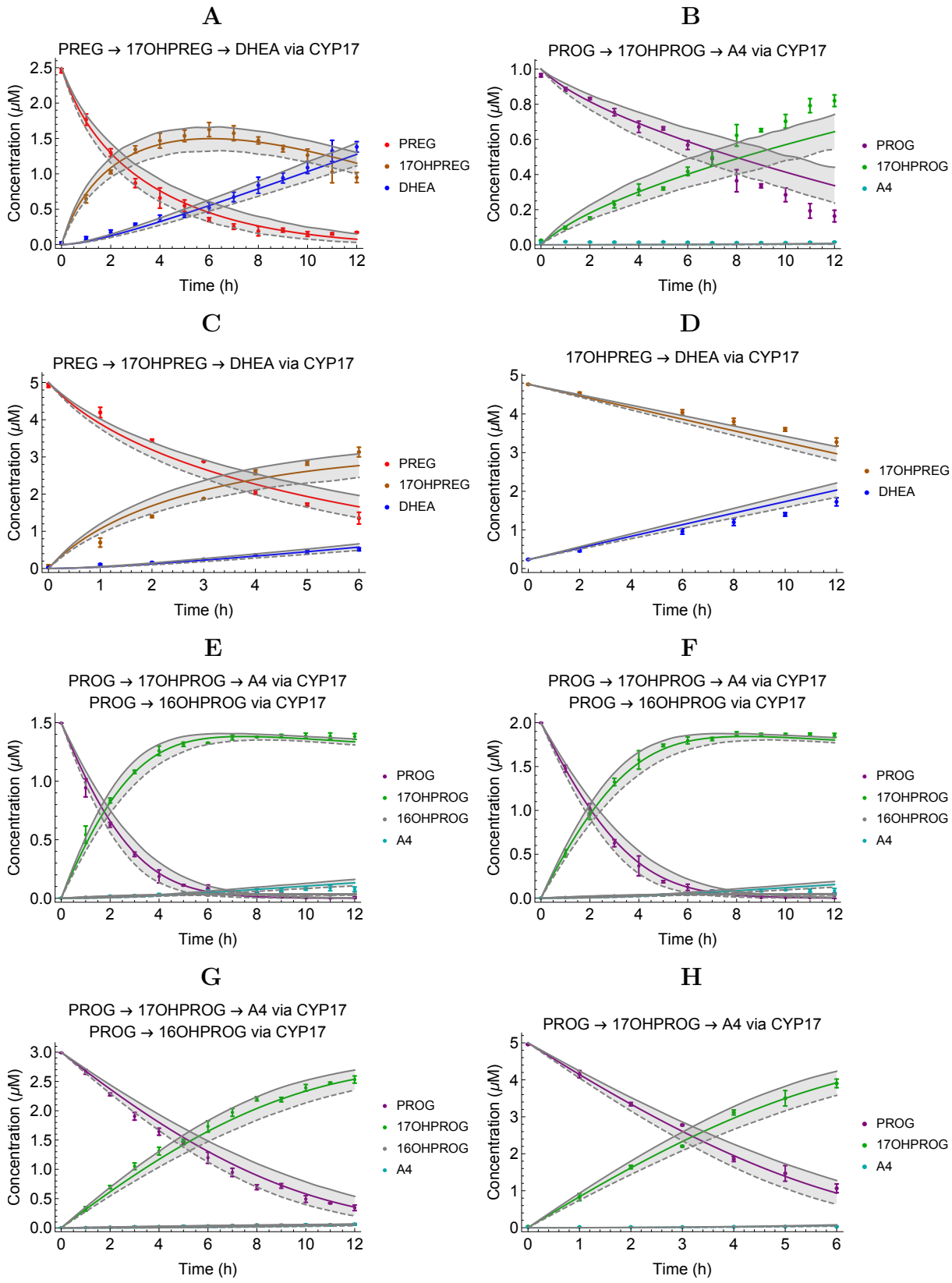


Figure 4.5: **Model fit to CYP17 data**. Model fit (solid lines) to the experimental data (points) used for the parameterisation of CYP17 parameters. The 95% confidence intervals are indicated by the grey areas.

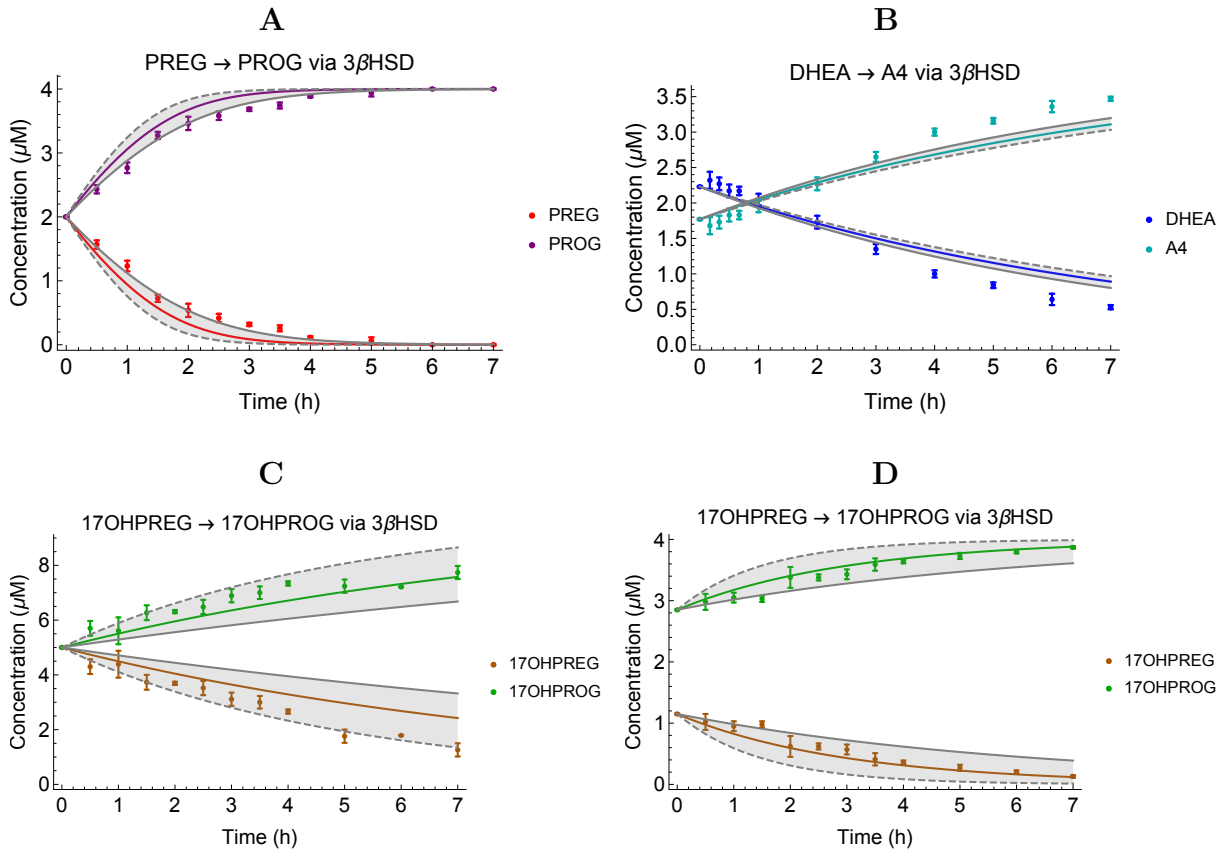


Figure 4.6: **Model validation of the 3βHSD parameters.** The experimental data (points) are shown along with the model prediction (solid lines).

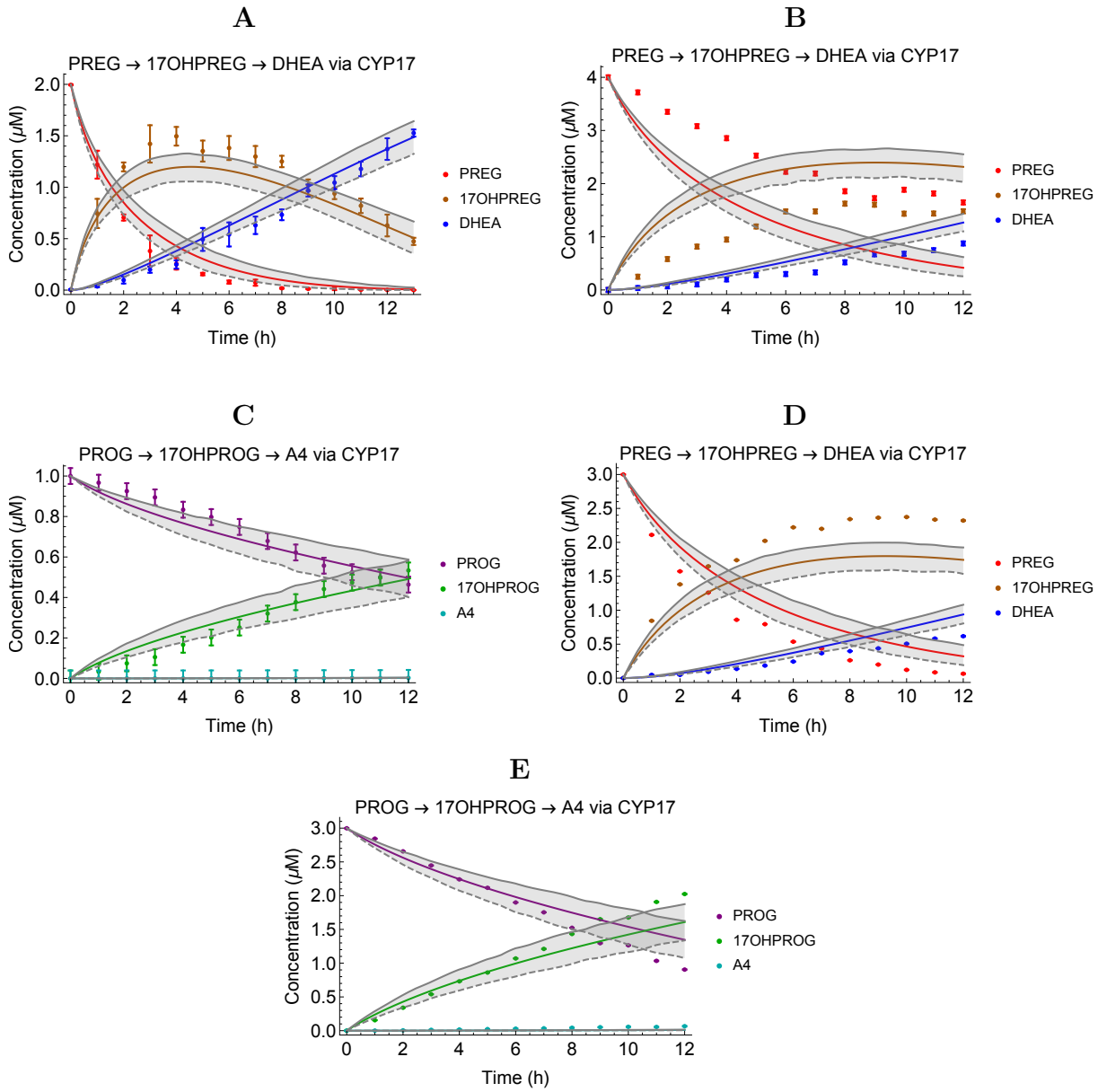


Figure 4.7: **Model validation of the CYP17 parameters.** The experimental data (points) are shown along with the model prediction (solid lines).

This grouping of the datasets into the respective construction and validation combinations returned the maximum number of identifiable parameter estimations, as discussed below.

4.1.2 Identifiability analysis

The method chosen to evaluate the identifiability of parameter estimations is the likelihood ratio test and profile likelihood plots as described on page 32. The functions `profileLikelihood` and `likelihoodRatioPlot` of the Mathematica add-on package `IdentifiabilityAnalysis` were used. An identifiability analysis was done on each of the parameter estimations of both the CYP17 and 3βHSD enzymes and confidence intervals for the newly determined parameter values were calculated.

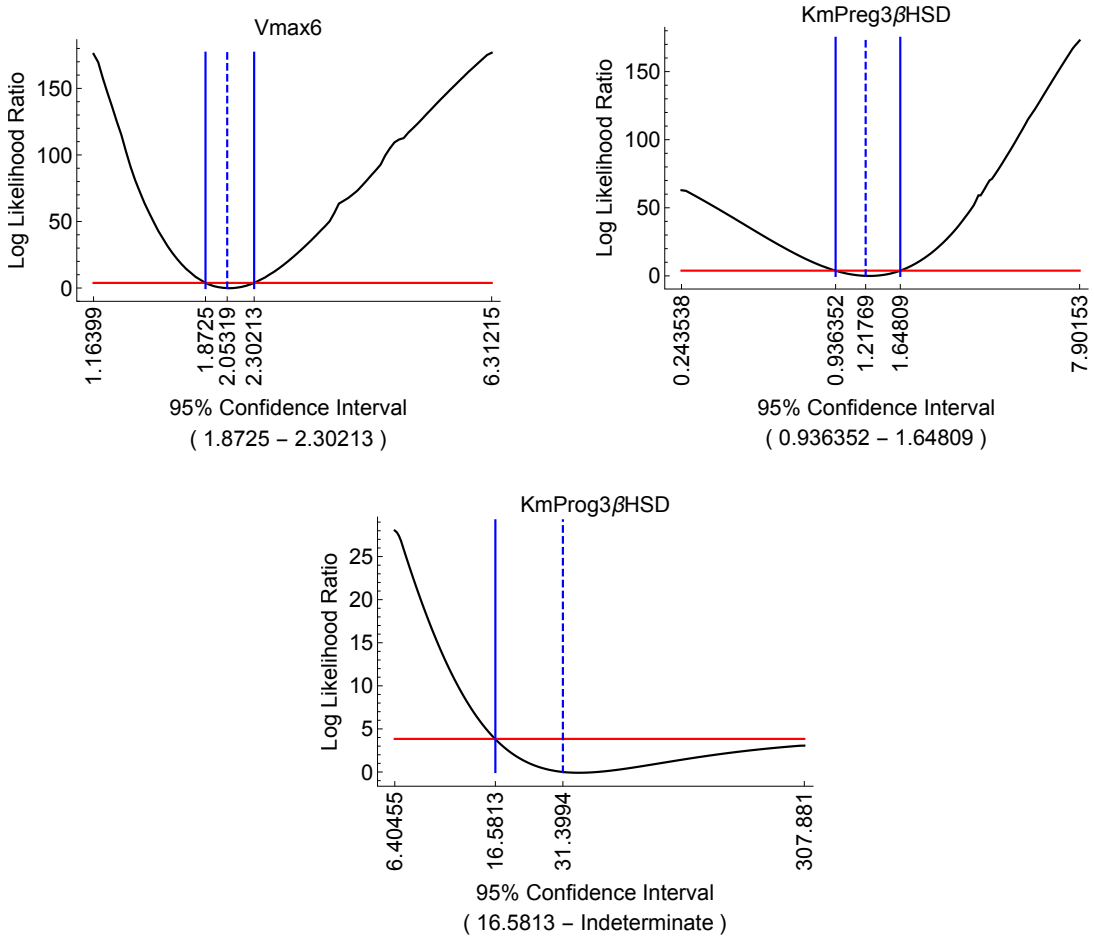


Figure 4.8: **Identifiability analysis of 3βHSD $Vmax_6$, $Km_{Preg3BHSD}$ and $Km_{Prog3BHSD}$ for the conversion of PREG to PROG.** Both $Vmax_6$ and $Km_{Preg3BHSD}$ are identifiable as these parameters are constrained within upper and lower bound confidence intervals. $Km_{Prog3BHSD}$ is not identifiable as this parameter is not constrained by an upper limit. The red lines indicate the 95% confidence cut-off value.

Figure 4.8 shows the profile likelihood plots for the three parameters $Vmax_6$, Km_{Preg} and Km_{Prog} for the conversion of PREG to PROG by 3βHSD, the only 3βHSD reaction with a non-identifiable parameter value. Both $Vmax_6$ and Km_{Preg} are identifiable as they are constrained within confidence intervals with upper and lower limits. Km_{Prog} is practically non-identifiable as it is only constrained by a lower limit confidence interval. This practical non-identifiability is a consequence of it being difficult to obtain experimentally high enough concentrations of product formation to fully parameterise and identify the product inhibition of 3βHSD by PROG. The practical non-identifiability can be considered in a quantitative sense, for instance if the 95% confidence is lowered to a 90% confidence, the parameter Km_{Prog} of 3βHSD would be identifiable. However, the 95% confidence level was used throughout, even though decreasing the confidence would in some cases have increased the number of identifiable parameters. All other

3βHSD parameter values were determined identifiably. The parameter values and their corresponding confidence intervals can be seen in Table 4.5. The profile likelihood plots of the other 3βHSD parameters can be seen in Appendix A on page 134.

Figure 4.9 shows the two profile likelihood plots for the parameters $Vmax_5$ and $Km_{17OHProg}$ for the conversion of 17-OHPROG to A4 by CYP17, the only CYP17 reaction with non-identifiable parameter values. These parameters are practically non-identifiable. The synthesis of A4 from 17-OHPROG via CYP17 is very low in the ovine species. The practical non-identifiability seen is due to the very low flux through this branch of the steroidogenic pathway [81]. All the other parameter values for the CYP17 enzyme were determined identifiably. These values and the confidence intervals can be seen in Table 4.6 and their profile likelihood plots can be seen in Appendix A on page 134.

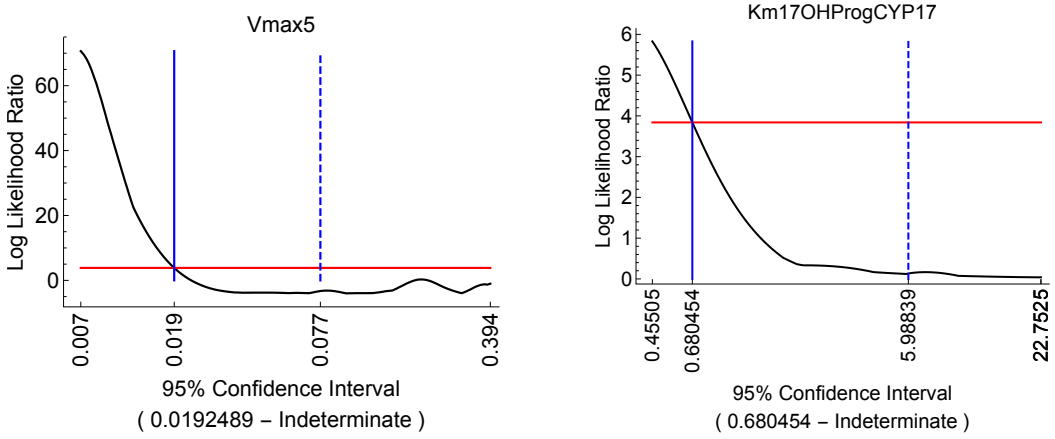


Figure 4.9: **Identifiability analysis of CYP17 $Vmax_5$ and $Km_{17OHProgCYP17}$ for the conversion of 17OH-PROG to A4.** Neither of these two parameters are identifiable as they are only constrained by lower limit confidence values. The red lines indicate the 95% confidence cut-off value.

Table 4.5: **3βHSD Parameter values and confidence intervals (CI)**

Parameter	Identifiability	Optimal value	Lower bound 95% CI	Upper bound 95% CI
$Vmax_6$	Identifiable	2.05319 (μM/h)	1.8725	2.30213
$Vmax_7$	Identifiable	2.82952 (μM/h)	2.01085	4.87385
$Vmax_8$	Identifiable	1.16183 (μM/h)	1.08423	1.24653
$Km_{Preg3BHSD}$	Identifiable	1.21769 (μM)	0.936352	1.64809
$Km_{Prog3BHSD}$	Not identifiable	31.39936 (μM)	16.5813	Indeterminate
$Km_{17OHPreg3BHSD}$	Identifiable	2.19637 (μM)	1.04871	5.60575
$Km_{17OHProg3BHSD}$	Identifiable	1.63111 (μM)	0.766163	3.90764
$Km_{DHEA3BHSD}$	Identifiable	0.23452 (μM)	0.15664	0.344214
$Km_{A43BHSD}$	Identifiable	0.217207 (μM)	0.135017	0.353842

Table 4.6: **CYP17 Parameter values and confidence intervals (CI)**

Parameter	Identifiability	Optimal value	Lower bound 95% CI	Upper bound 95% CI
$Vmax_1$	Identifiable	1.67292 (μM/h)	1.36737	2.11747
$Vmax_2$	Identifiable	0.152892 (μM/h)	0.12613	0.175953
$Vmax_3$	Identifiable	0.798223 (μM/h)	0.722805	0.899345
$Vmax_4$	Identifiable	0.0171024 (μM/h)	0.0123062	0.0215451
$Vmax_5$	Not identifiable	0.0777709 (μM/h)	0.0192489	Indeterminate
$Km_{PregCYP17}$	Identifiable	0.550419 (μM)	0.298424	0.883136
$Km_{ProgCYP17}$	Identifiable	0.884594 (μM)	0.263981	1.09416
$Km_{17OHPregCYP17}$	Identifiable	0.17159 (μM)	0.0467976	0.213687
$Km_{17OHProgCYP17}$	Not identifiable	5.98839 (μM)	0.680454	Indeterminate

4.2 Establishing the ovine model as a control model: Comparison with results in literature

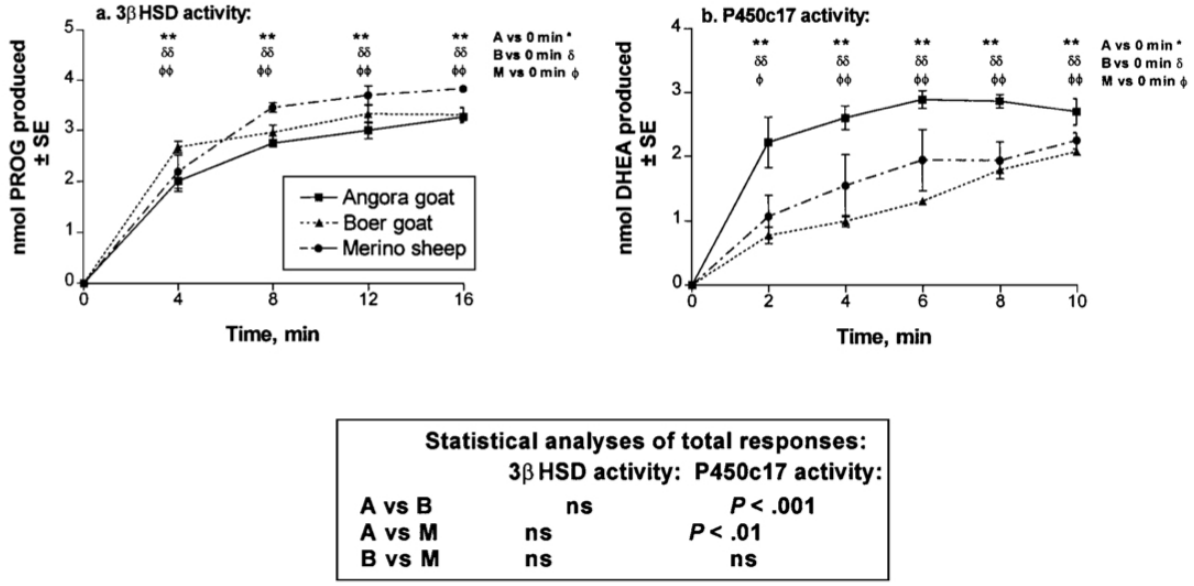
In this section the newly parameterised ovine model is validated against results in literature. The model is first validated by predicting the experimental results of a study by Engelbrecht and Swart [28]. Thereafter the model is used to simulate the theoretical results of a study by Nguyen *et al* [81]. These validations of the model against literature results establishes the model as valid reference model for steroid synthesis.

4.2.1 Simulating experimental results for the ovine species

In section 4.1 the model is parameterised with experimental data collected in a reconstituted system with the enzymes CYP17 and 3βHSD expressed in transfected COS-1 cells. The V_{max} and K_m values for the substrates and products of the conversion of one steroid to the next were determined. In the model these V_{max} values are the total enzyme concentration times the catalytic rate constant of the enzyme for a specific conversion [95]. With variation in the transfection efficiency in *in vitro* studies and unknown expression levels of the enzyme in *in vivo* studies, it is not always possible to determine the true V_{max} values. In this section the V_{max} values are adjusted to account for the unknown expression of the enzymes in the published study. This is done by adding a variable multiplier to the rate equation which is then fitted for. This is a necessary step when comparing model predictions to results of other independent studies when the enzyme concentration is unknown.

The model is validated by simulating the outcome of experiments not conducted with transfected cells, but rather the adrenal microsomes from harvested animal tissue. By adjusting the V_{max} values of the model to account for the enzyme expressions in

the published study, the ratio of the enzyme activities are also adjusted to represent physiological enzyme expression. This validation tests whether the model can predict the steroidogenesis in ovine adrenal microsomes. In the study by Engelbrecht and Swart, the adrenal steroidogenesis of the Angora goat was compared to that of the Merino sheep and Boer goat. Both the Boer goat and Merino sheep are considered to be hardy species which do not suffer from the hypocortisolism seen in the Angora goats. If the model can accurately predict the outcome of the ovine results in the study, it can be used as a reference model for steroidogenic function.



Statistical analyses of total responses:	3β HSD activity:	P450c17 activity:
A vs B	ns	$P < .001$
A vs M	ns	$P < .01$
B vs M	ns	ns

Figure 4.10: **3βHSD and CYP17 activity in Angora goat, Boer goat and Merino sheep adrenal microsomes**. (a) Comparison of 3βHSD activity in Angora goats, Boer goat and Merino sheep adrenal microsomes by measuring the conversion of 10μM PREG to PROG. (b) Comparison of CYP17 activity in Angora goats, Boer goat and Merino sheep adrenal microsomes by measuring the conversion of 10μM PREG to 17-OHPREG and DHEA. The figure axis is labeled only as the the production of DHEA, but this should be labeled as the the production of DHEA and 17-OHPREG in total (confirmed by author). Figure taken directly from [28].

In the study by Engelbrecht and Swart, the microsomes from the adrenal glands of freshly slaughtered animals were used to collect steroidogenic data [28]. Figure 4(a) in the study, shown here in Fig 4.10(a), compares the 3βHSD activity of the Angora goat, Boer goat and Merino sheep. This was done by comparing the conversion of 10μM PREG to PROG between the species. Figure 4(b) in the study, shown here in Fig 4.10(b), compares the CYP17 activity of these three species by measuring the conversion of 10μM PREG to 17-OHPREG and DHEA. The activity of each of the enzymes were studied independently of the other by adding the appropriate cofactors to the adrenal microsome reaction mixture. NADPH was added to measure the activity of CYP17 for

the conversion of PREG to 17-OHPREG and DHEA and the cofactor NAD^+ was added to measure the activity of 3βHSD for the conversion of PREG to PROG [28].

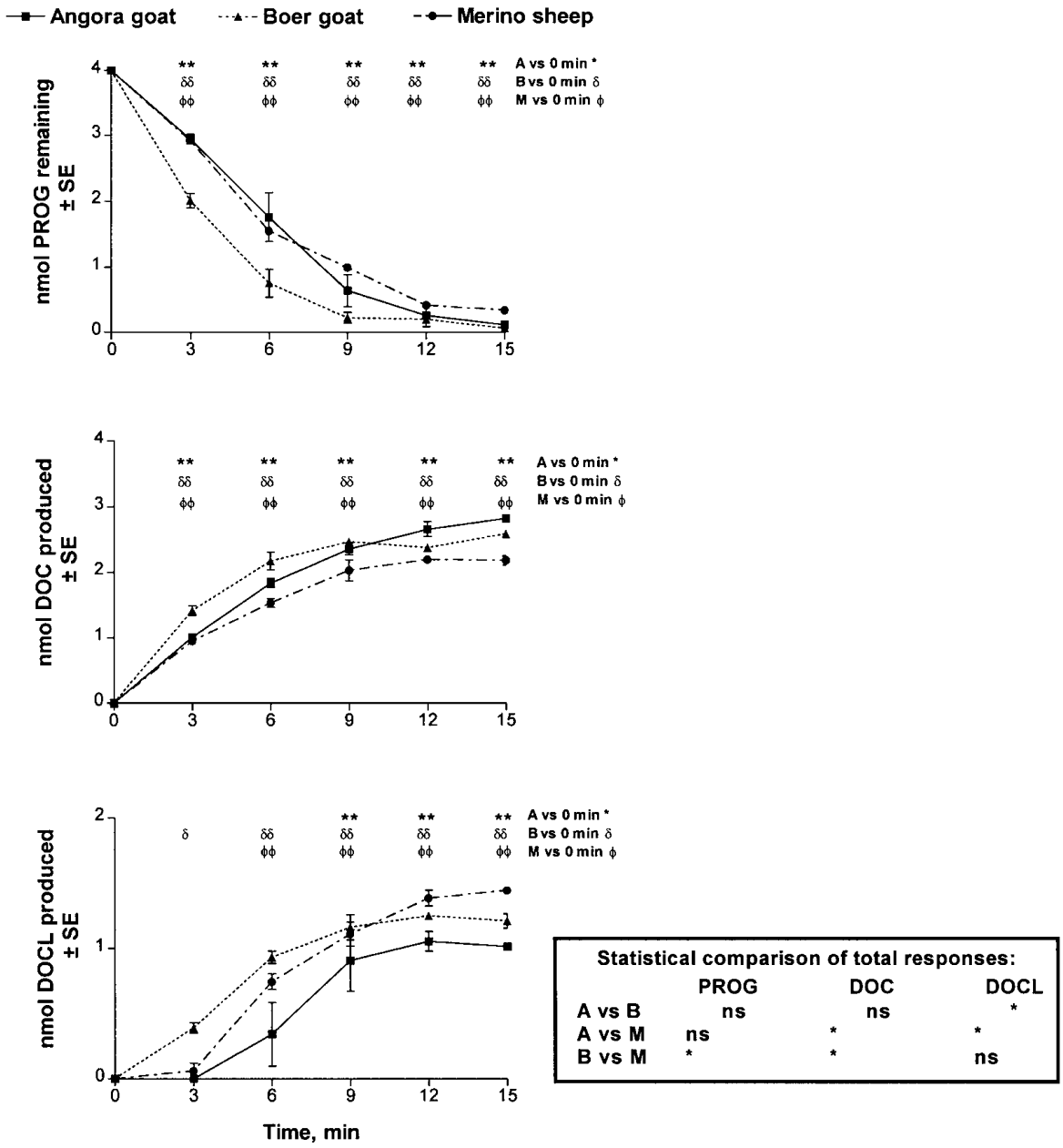


Figure 4.11: **Time course data of 10μM PROG metabolism by Angora goat, Boer goat and Merino sheep microsomes.** Figure adapted from Fig. 2 in [28].

Figure 2 in the study, shown here in Fig. 4.11, compares the conversion of PROG to 17-OHPROG, DOC (Deoxycorticosterone), DOCL (Deoxycortisol), and A4 between the three species. This experiment was also conducted with the microsomes from the adrenal glands of the three different species. The CYP17 enzyme converts PROG

to 17-OHPROG and in turn converts 17-OHPROG to A4, while the CYP21 enzyme converts PROG and 17-OHPROG to DOC and DOCL respectively. The syntheses of 17-OHPROG and A4 were insignificant and therefore not published in their study. The experiment was initiated with 10μM PROG.

All of the above mentioned experiments were normalised with the adrenal P450 content (0.32 mM P450). The P450 concentration was determined by spectrophotometric assays. The total protein was not recorded.

To validate the newly parameterised ovine model, simulations of these experimental results were created. The expression levels of the CYP17, 3βHSD, and CYP21 enzymes in the adrenal glands of the sheep, used in these experiments, are unknown. To account for the unknown expression levels in this study, three multipliers were fitted for, α, β, and γ. These multipliers represent the expression of the enzymes 3βHSD, CYP17, and CYP21 respectively. Although fitted for, these multipliers do adhere to constraints. The values of α and β are firstly fitted for by using the two time courses of Merino sheep data in Figure 4.10. In all subsequent analysis, the values of α and β are fixed, and not fitted for again. When the multiplier γ is fitted for, with some of the time course data of Figure 4.11, the values of the other variable multipliers are fixed to their previously determined values. During the fitting for the variable multipliers and all further model simulations, the values of the 18 parameters, present in the model, were fixed to the values in Tables 4.5 and 4.6, as previously determined.

The α and β multipliers were fitted for by using the Merino sheep data points of Fig. 4.10(a) and Fig. 4.10(b) respectively. Engelbrecht and Swart studied the kinetics of 3βHSD and CYP17 independently by adding the appropriate cofactors. Fig. 4.10(a) show only the kinetics of 3βHSD and Fig. 4.10(b) show only the kinetics of CYP17. The ovine model was simplified to mimic the experimental conditions of these two figures by setting the rate equations of the pathway branches, not present in the experiment, equal to zero. The resulting rate equations for the fitting of α and β are equations 4.4 and 4.6 respectively, while the ODEs are equations 4.5 and 4.7. To fit for α and β an objective function in Mathematica finds the optimal solution for these multipliers by minimising the distance between the time course data for the Merino sheep of Fig. 4.10(a) and 4.10(b) and the model simulation. The model simulation of the time course data in Fig. 4.10(a) and 4.10(b) can be seen in Fig. 4.12. The values identified as the optimal and identifiable solutions for the parameters, shown in Tables 4.5 and 4.6, were used as the values for the parameters in equations 4.4 to 4.10. The fitted values for α and β are 43.86 and 56.96 respectively. These two values are used in all subsequent analysis to adjust for the relative expression of 3βHSD and CYP17 in the ovine species.

Rate equation for determining the activity of 3βHSD:

$$v_1 = \frac{\alpha \cdot Vmax_6\left(\frac{[PREG]}{Km_{Preg3\beta HSD}}\right)}{1 + \frac{[PREG]}{Km_{Preg3\beta HSD}} + \frac{[PROG]}{Km_{Prog3\beta HSD}}} \tag{4.4}$$

ODEs for determining the activity of 3βHSD:

$$\begin{aligned} \frac{d}{dt}[PREG] &= -v_1 \\ \frac{d}{dt}[PROG] &= v_1 \end{aligned} \tag{4.5}$$

Rate equations for determining the activity of CYP17:

$$\begin{aligned} v_2 &= \frac{\beta \cdot Vmax_1\left(\frac{[PREG]}{Km_{PregCYP17}}\right)}{1 + \frac{[PREG]}{Km_{PregCYP17}} + \frac{[17OHPREG]}{Km_{17OHPregCYP17}}} \\ v_3 &= \frac{\beta \cdot Vmax_2\left(\frac{[17OHPREG]}{Km_{17OHPregCYP17}}\right)}{1 + \frac{[17OHPREG]}{Km_{17OHPregCYP17}}} \end{aligned} \tag{4.6}$$

ODEs for determining the activity of CYP17:

$$\begin{aligned} \frac{d}{dt}[PREG] &= -v_2 \\ \frac{d}{dt}[17OHPREG] &= v_2 - v_3 \\ \frac{d}{dt}[DHEA] &= v_3 \end{aligned} \tag{4.7}$$

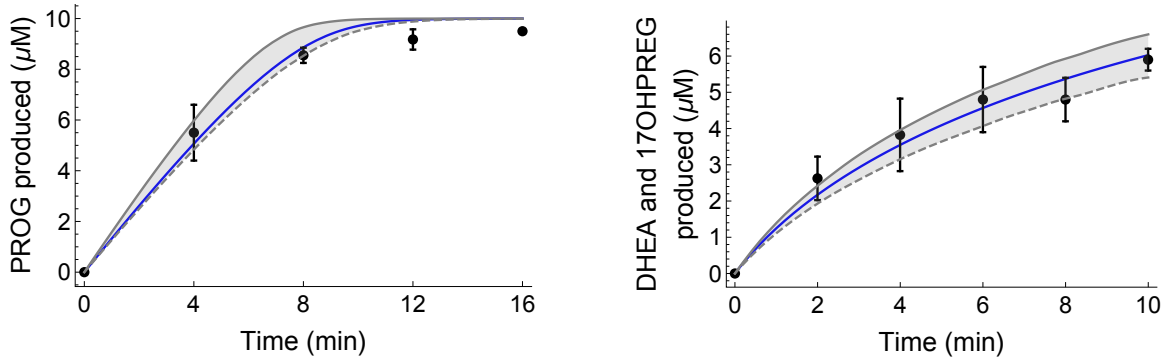


Figure 4.12: **Model simulation of Fig 4.10(a) and 4.10(b) where the multipliers α and β are fitted for.** The model simulation is depicted as the solid line and the time course data for the Merino sheep from the Fig. 4.10 are shown as the black data points. Data is converted to concentration data (μM). The 95% confidence intervals of the model simulation is shown by the grey areas.

The multiplier γ represents the expression of the CYP21 enzyme in the experiment shown in Fig. 4.11. CYP21 converts PROG and 17-OHPROG to DOC and DOCL respectively. These two branches of the pathway have not yet been included in our model, therefore mass action kinetics were used to account for the synthesis of DOC and DOCL. In the study by Engelbrecht and Swart the initial rate of DOC production was found to be 0.25 $nmol{\cdot}mL^{-1}{\cdot}min^{-1}$ and that of DOCL was found to be 0.13 $nmol{\cdot}mL^{-1}{\cdot}min^{-1}$ [28]. Only one variable multiplier is fitted for for both branches by including the ratio of DOC formation relative to DOCL formation. Therefore the formation of DOC is equivalent to the constant value of γ times the concentration of PROG and the formation of DOCL is equivalent to γ times the concentration of 17-OHPROG and the ratio of 0.13/0.25, see equations 4.9 and 4.10.

The value of 56.96 for the multiplier β was substituted into equation 4.8. With the value for β determined for one experiment, we simulated the results of another experiment in the study.

The multiplier γ was fitted for by using the Merino sheep data points of Fig. 4.11. The model was simplified to mimic the experimental conditions of this figure by setting the rate equations of the pathway branches, not used in the experiment, equal to zero. Although the conversion of PROG to 16-OHPROG is included in this model for the ovine species, the levels of 16-OHPROG accumulation is low enough that the inclusion or exclusion of 16-OHPROG formation has no effect on overall PROG consumption and 17-OHPROG formation. It has therefore been excluded from these model simulations. The resulting rate equations are 4.8 to 4.10 and the ODEs are 4.11. These equations are fitted to the experimental data in Fig. 4.11 and the model simulation of the experimental

time course data can be seen in Fig. 4.13. From a formal mathematical perspective the model could be rejected because the experimental data points fall outside the 95% confidence intervals. However, from a biological perspective, the model was constructed with data from a different system (transfected cells) than the microsome system of which the dynamics is predicted in Fig. 4.13. The model, being able to predict the outcome of the microsome data to the level of accuracy seen in Fig. 4.13, is a useful tool for studying the steroidogenic dynamics in this system.

Rate equations for determining the expression of CYP21:

$$v_4 = \frac{\beta \cdot Vmax_3\left(\frac{[PROG]}{Km_{ProgCYP17}}\right)}{1 + \frac{[PROG]}{Km_{ProgCYP17}} + \frac{[17OHPROG]}{Km_{17OHProgCYP17}}} \tag{4.8}$$

$$v_5 = \gamma \cdot [PROG] \tag{4.9}$$

$$v_6 = \gamma \cdot \frac{0.13}{0.25} \cdot [17OHPROG] \tag{4.10}$$

ODEs for PROG conversion to 17-OHPROG, DOC, and DOCL:

$$\begin{aligned} \frac{d}{dt}[PROG] &= -v_4 - v_5 \\ \frac{d}{dt}[17OHPROG] &= v_4 - v_6 \\ \frac{d}{dt}[DOC] &= v_5 \\ \frac{d}{dt}[DOCL] &= v_6 \end{aligned} \tag{4.11}$$

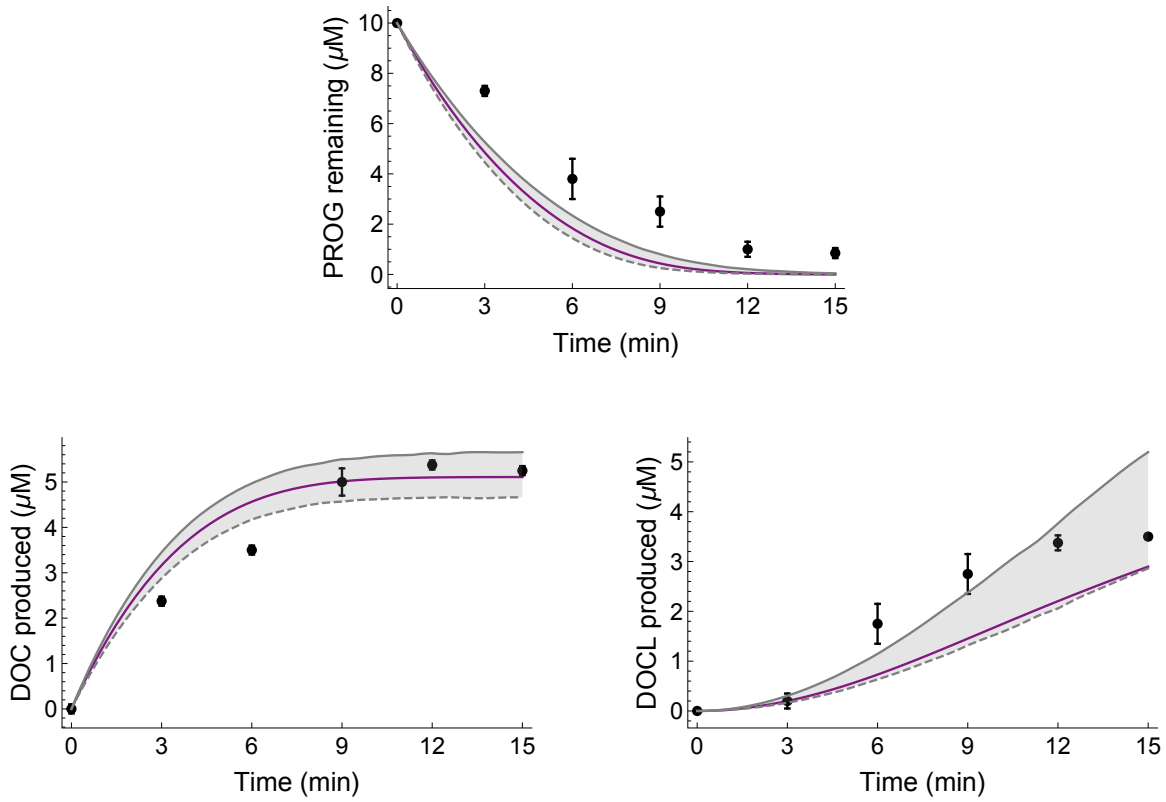


Figure 4.13: **Model simulation of Fig. 4.11 where the multiplier γ is fitted for.** The model simulation is depicted as the solid line and the time course data for the Merino sheep from the Fig. 4.11 is shown as the black data points. The 95% confidence intervals of the model simulation is shown by the grey areas.

With the values of α, β and γ determined, the results of another experiment, published in the same study by Engelbrecht and Swart, were predicted [28]. The experimental results for the percentage metabolite distribution of 10 μM PREG metabolism in the adrenal microsomes of the Merino sheep (Table 1 in the study [28]) are shown here in Table 4.7. This table also contains the model prediction of these experimental results; the percentage conversion of 10μM PREG to DOC, DOCL, 17-OHPREG, DHEA, and A4 with the simulation allowed to run to completion. The model was given the initial metabolite concentrations, used to conduct the experiment, and predicted the intermediate and product synthesis. When comparing the model simulation with the Merino sheep data, the synthesis of DOC and DOCL, and DHEA and 17-OHPREG are fairly similar. The ovine model predicted the synthesis of DOC and DOCL to be 80.65% of the total metabolite formation, while 0% DHEA and 17-OHPREG were formed. The published result was a value of 82.03% for DOC and DOCL synthesis, while DHEA and 17-OHPREG synthesis were also 0% in the sheep microsomes. The model predicted the synthesis of A4 being 18.39% of the total metabolite distribution while 0% of the PREG remained, it is however uncertain what the A4 synthesis and remaining PREG were in the ovine microsomes as the values were not published.

Table 4.7: **Percentage metabolite distribution of 10 μM PREG metabolism in the adrenal microsomes of the Merino sheep** (taken from [28]) and the model prediction of the percentage metabolite distribution.

	% DOC & DOCL	% DHEA & 17-OHPREG	% A4	% PREG
Merino sheep microsomes	82.03	0.00	-	-
Model prediction	80.65	0.00	18.39	0.00

With these same values of α, β and γ for the relative expression of CYP17, 3βHSD, and CYP21, the model was used to simulate the outcome of another independent experiment completed by Engelbrecht [26]. The effect of PREG on the production of PROG was studied. The production of 17-OHPROG, DOC, and DOCL was recorded in the presence and in the absence of 10 μM PREG. Similarly to the experimental conditions of the study by Engelbrecht and Swart [28], these experiments were also conducted with the microsomes from the adrenal glands of the Angora goats, Merino sheep, and Boer goats. This experiment was normalised with the value of 0.22 mM P450. With the multipliers α, β, and γ fitted to experimental data normalised with the value of 0.32 mM P450, the multipliers were scaled with the ratio of 0.22/0.32 for the simulation of these results. The experimental Merino sheep results of this experiment are summarised in Table 4.8 as well as the model prediction of this experiment for the Merino sheep.

Table 4.8: **Products of 10 μM PROG metabolism in Merino sheep adrenal microsomes in the presence and absence of 10 μM PREG in the incubation mixture** (taken from [28]) and the model prediction thereof.

	%PROG remaining		%17OHPROG & DOC		%DOCL	
	No PREG	10μM PREG	No PREG	10μM PREG	No PREG	10μM PREG
Merino sheep	4	48	64	41	31	10
Model prediction	1.97	15.65	79.78	67.10	17.33	6.97

In their study it was found that all three species metabolised less PROG in the presence of PREG. In the Merino sheep the amount of unmetabolised PROG increased 12 times, our model showed nearly a 10 times increase in unmetabolised PROG. It was also found that the Merino sheep produces less DOCL, 17-OHPROG and DOC in the presence of PREG than in the absence thereof. Their study showed a 36% decrease in 17-OHPROG

and DOC production, while our results showed a 16% decreased in the production of these metabolites. The production of DOCL decreased with 68%, with our model showing a decrease of 60% [26]. Our model is fairly successful in predicting the same trends in steroid synthesis, in the presence and absence of PREG, as seen in the experimental study.

The results of this section indicate that the ovine model can adequately simulate the results of experimental data generated with ovine adrenal microsomes, shown in Table 4.7. The ovine model could also qualitatively predict the relative difference in steroid synthesis in the presence and the absence of PREG (Figure 4.8).

4.2.2 Simulating theoretical results for the ovine species

In this section the ovine model is validated against a theoretical model in literature. The model by Nguyen *et al* [81] describes steroidogenesis in the human, non-human primates, ovine, and bovine species. As an additional validation of our ovine model the results published in the Nguyen study is recreated.

As mentioned in the previous chapter, a paradoxical increase of E2 is seen in rhesus monkeys and sheep with inhibition of 3βHSD [81]. A4 is a precursor for E2. Figure 4.14 shows the Δ^4 and Δ^5 branches of the steroid hormone biosynthesis pathway as modelled by Nguyen *et al* [81]. This model of the pathway does not include the CYP17 reaction where 17-OHPROG is converted to A4 as this reaction does not take place in all mammals. As such, 3βHSD is the only enzyme facilitating the synthesis of A4. It is therefore counterintuitive to see an increase in E2 synthesis when 3βHSD is inhibited.

The Nguyen model showed that with negligible conversion of 17-OHPROG into A4 via CYP17, A4 production can be increased with decreased 3βHSD activity. In their study it is shown that a constant decrease in 3βHSD activity causes a constant decrease in the production of PROG and 17-OHPROG. However, with decreased 3βHSD activity A4 production is increased up to a certain point, but further decrease in 3βHSD activity leads to a sharp decrease in A4 synthesis. Sensitivity analysis showed that these results are only dependent on V_{max} values. Their study also showed that a decrease in the affinity of CYP17 for PREG (increasing the Km value 10-fold) showed similar qualitative results [81]. This phenomenon in A4 production is also seen in *in vivo* systems. Rhesus monkeys produce more E2 during the mid-luteal phase of the menstrual cycle. In another example lutectomised and ovarectomised 90-day pregnant ewes produce more E2 when 3βHSD is inhibited by trilostane [24, 81, 122].

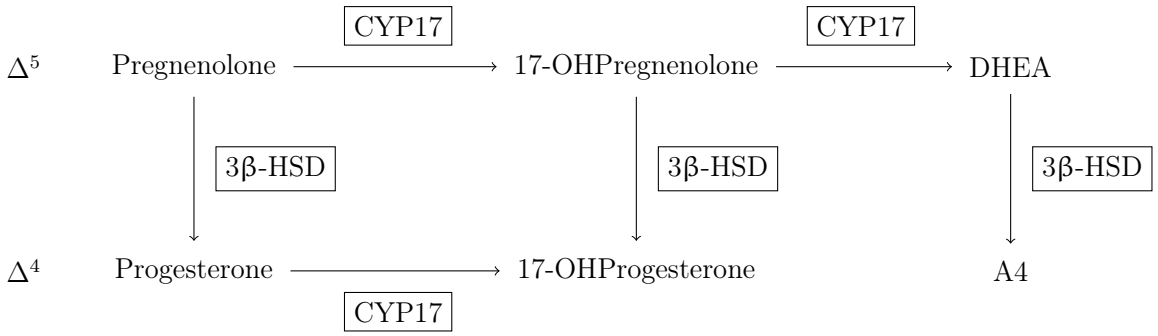


Figure 4.14: **Schematic representation of the Δ^4 and Δ^5 branches of the steroid hormone biosynthesis pathway as modelled by Nguyen *et al* [81].**

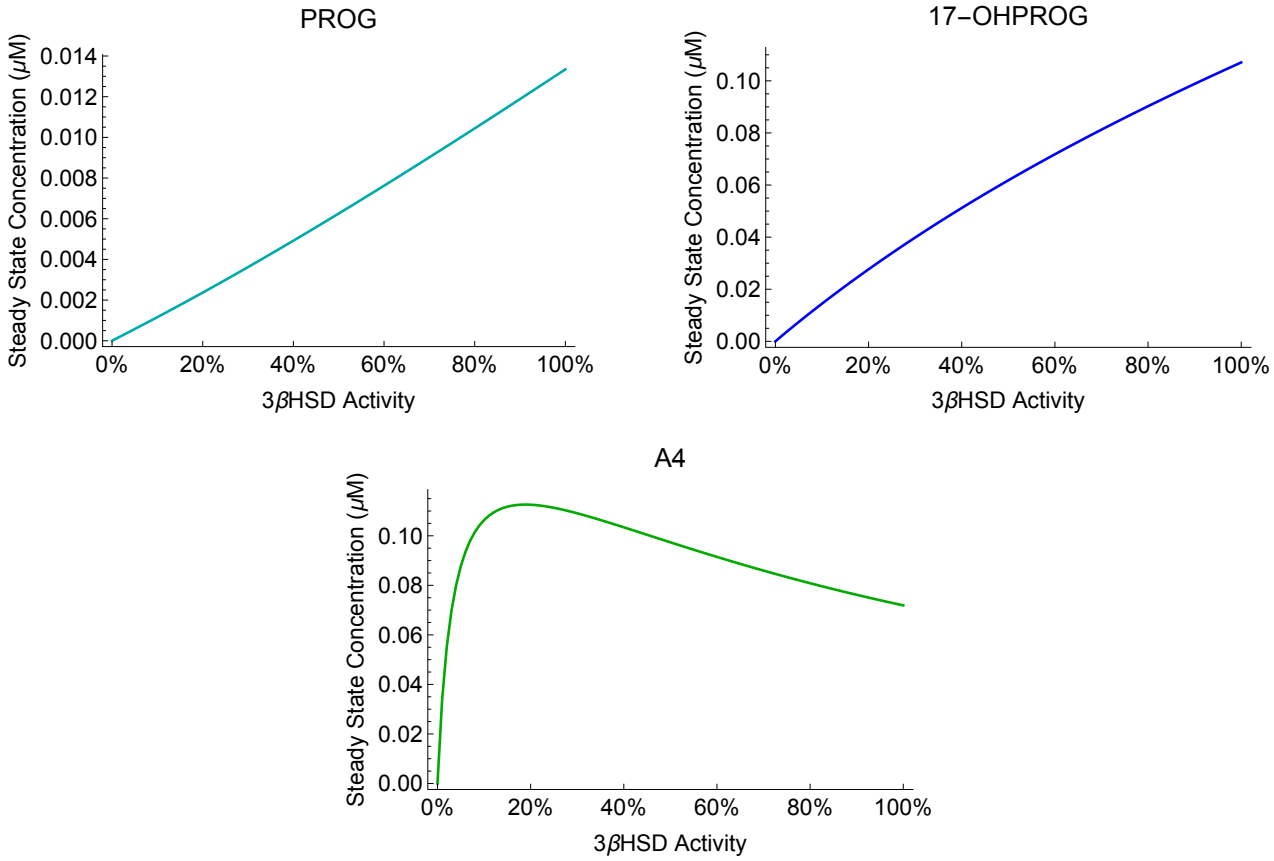


Figure 4.15: **Decreased synthesis of PROG and 17-OHPROG, and increased synthesis of A4 seen with decreased 3βHSD activity.** The concentration of A4 rapidly decreases with further decrease in 3βHSD activity beyond a certain point.

This paradoxical increase in E2 seen with a decrease in 3βHSD activity can also be simulated with our model (Fig. 4.15). A4 is a precursor for E2 and with a decrease in the activity of 3βHSD, an increase in A4 synthesis can be simulated. A decrease in the activity of 3βHSD beyond a certain point then also leads to a sharp decrease in A4 synthesis. However, a steady decrease in enzyme activity leads to constant decreased synthesis of 17-OHPROG and PROG. These simulations correspond to the behaviour

of the model by Nguyen *et al* [81]. Sensitivity analysis on our model also showed that similar qualitative results are returned with variation in all other parameter values, apart from the activity of 3βHSD. The ratio of the activities of CYP17 and 3βHSD also does not have any influence i.e. a constant increase in the activity of CYP17 does not return the same results. This phenomenon is only seen with a decrease in the activity of 3βHSD.

The Nguyen model also showed that increased PREG supply initially increases the production of all metabolites. Their steady state analysis showed that increasing the PREG supply rate beyond a certain level decreases the production of A4, then DHEA, and finally 17-OHPROG, but continually increases PROG production. Sensitivity analysis showed that these results remain unchanged with variation in all other parameter values, apart from the PREG supply [79].

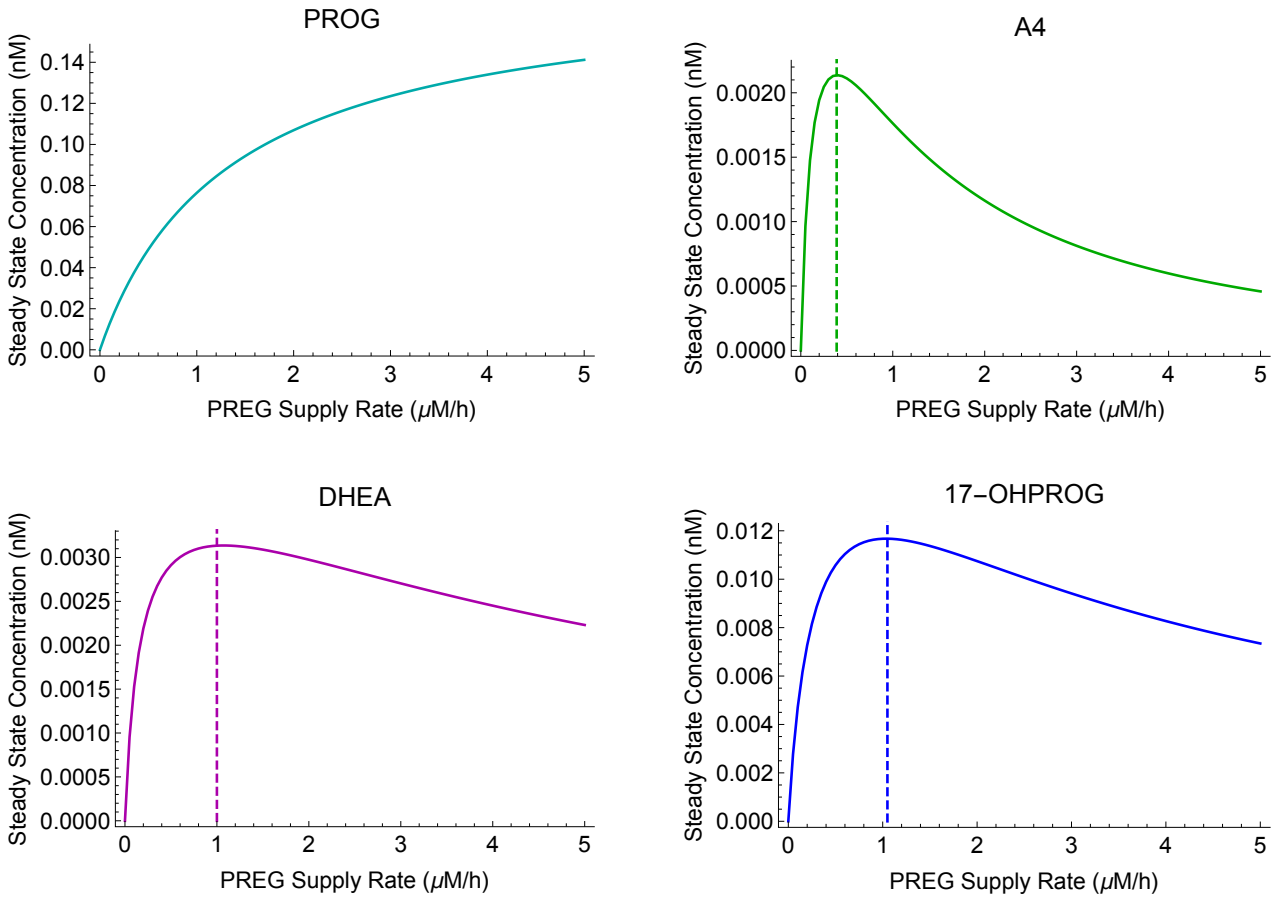


Figure 4.16: **Changes in steady state concentrations of PROG, A4, DHEA, and 17-OHPROG with varying PREG supply rate.** The PREG supply ranges from 0 to 5 μM/h. The steady state concentration of A4 decreases first when the PREG supply rate nearly reaches 0.4 μM/h. The steady state concentration of DHEA decreases next, once the PREG supply rate reaches 1.0 μM/h, while the concentration of 17-OHPROG decreases last, once the PREG supply is just above 1.0 μM/h

These results could also be simulated with our model. Figure 4.16 shows that an increase in the supply rate of PREG causes a continuous increase in the PROG steady state concentration. Figure 4.16 also shows an initial increase in A4, DHEA, and 17-OHPROG production with an increase in the PREG supply rate, but then a decrease in metabolite production is seen with further increase in PREG supply. Similar to the results obtained with the Nguyen model, our model shows a suppression of A4 synthesis first, then the suppression of DHEA, and finally the production of 17-OHPROG is suppressed with increase in PREG supply beyond a certain rate.

The results in this section and the previous section show that the newly parameterised ovine model could fairly accurately simulate the results of two independent studies in literature. Thus the ovine model can be used further as a reference model for healthy steroidogenesis in the process of studying hypocortisolism in Angora goats.

4.3 Development of the Angora model

In this section we use the ovine model, as parameterised and validated in the previous sections, to create a model that represents the steroidogenesis in the Angora goat. The Angora model is parameterised by combining the ovine model with the altered enzyme activities of the Angora goat (taken from the study by Engelbrecht and Swart [28]), whereafter the model is validated against an independent study in literature.

Engelbrecht and Swart found that the activity of the 3βHSD enzyme was 0.53 $nmol.mL^{-1}.min^{-1}$ for the conversion of PREG to PROG and the activity of CYP17 was 0.5 $nmol.mL^{-1}.min^{-1}$ for the conversion of PREG to DHEA in the Merino sheep. These activities were measured using the microsomal subcellular fractions of the homogenised adrenal microsomes of seven Merino sheep ewes, resuspended in 100 mM potassium phosphate buffer. The activity of these two enzymes are similar in the ovine species, but in the Angora goat the activity of CYP17 is higher than that of 3βHSD with 3βHSD having an activity of 0.5 $nmol.mL^{-1}.min^{-1}$ while CYP17 has an activity of 1.13 $nmol.mL^{-1}.min^{-1}$ [28]. These activities were measured using the microsomal subcellular fractions of eight Angora goat ewes, resuspended in 100 mM potassium phosphate buffer. The activity of CYP17 is more than twice (2.26 times) that of 3βHSD in the Angora goat. Increased CYP17 function may lead to reduced production of PROG and 17-OHPROG, precursors for the production of cortisol. The increased CYP17 activity can also lead to an increase in DHEA and A4 production.

To investigate the effects of this increased CYP17 activity seen in the Angora goats, a model is developed that represents the steroidogenesis in Angora goats. One should be

able to compare the dynamics of this Angora model with that of the reference ovine model, and the hypothesis is that the difference is due to an increase in CYP17 activity only. Therefore the model was not reparameterised with Angora goat experimental data, to exclude the effects that other parameter values may have on the outcome. The parameter values of the ovine species were used, and the increased activity of the CYP17 enzyme in the Angora goat was included. To create the Angora model, the CYP17 activity of the ovine model was set to be 2.26 times faster than the 3βHSD activity. The model is from here on further referred to as the "Angora model", whereas the original model is still referred to as the "ovine model". The Angora model is first compared to Angora time course data in literature to determine if it does resemble steroidogenesis in the Angora goat. Thereafter a comparison of the two models is made by considering the metabolic control analysis in both systems as well as the results of steady state analysis.

4.3.1 Validation of Angora model against results in literature

In a study by Storbeck *et al* [105] a single time course for the metabolism of 1 μM PREG to 17-OHPREG, DHEA, PROG, 17-OHPROG, 16-OHPROG, and A4 in reconstituted cells for the Angora goat is found. The Angora model is validated by predicting the outcome of this dataset. In addition to adjusting the activity levels of CYP17 and 3βHSD to that of the Angora goat, a new variable multiplier, δ, is fitted for in the same way as the other variable multipliers as described on page 58. This multiplier adjusts for the expression of both CYP17 and 3βHSD (i.e. only the total protein) in the experiment by Storbeck *et al* [105]. The value of the multiplier, δ , fitted for is 0.004783.

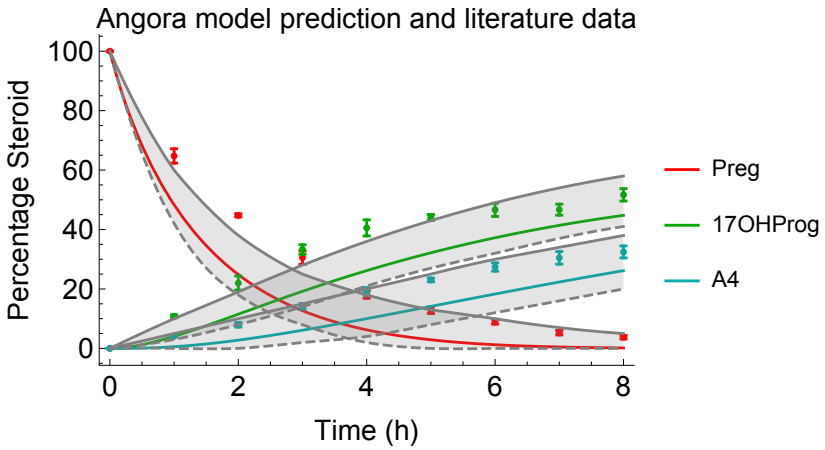


Figure 4.17: **Time course data and model prediction of the consumption of 1 μM PREG, and the synthesis of 17-OHPROG and A4 in the Angora goat.** Time course data taken from [105]. The 95% confidence intervals of the model simulation is shown by the grey areas.

Figure 4.17 contains the time course data for the consumption of 1 μM PREG, and

the synthesis of 17-OHPROG and A4 in the Angora goat as published in [105]. This figure also shows the predicted outcome of PREG consumption, and 17-OHPROG and A4 synthesis in this experiment by the Angora model. With the Angora model prediction of the experiment being very similar to the experimental time course data it is fair to say that the Angora model, as constructed by simply increasing the CYP17 activity of the ovine model, is representative of steroidogenesis in the Angora goat.

4.4 Hypocortisolism - comparison of the ovine and Angora models

The following section shows a comparison of the two models in terms of steady state and metabolic control analysis and how the difference in the kinetics of steroidogenesis relates to the symptoms of hypocortisolism.

The steady state concentrations of all steroids of the ovine model and the Angora model are compared first, as can be seen in Figure 4.18. For this specific steady state simulation, both models had a constant supply of 0.01 μM/h PREG and the initial concentrations of all other metabolites were zero. This simulation also includes the removal of A4 at a constant rate of k times the metabolite concentration where $k = 0.1$ h^{-1}. The model at this point in time does not include the conversion of A4 into subsequent metabolites, therefore this removal of A4 is included in the steady state analysis. This is similar to the steady state analysis method of the steroidogenic model seen in the study by Nguyen *et al* [81]. The ovine model, which represents healthy steroidogenesis, is shown in green, while the Angora model, with altered adrenal function, is shown in purple (Fig. 4.18). The Angora model has higher steady state concentrations for DHEA and A4, while it has lower steady state concentrations for 17-OHPROG than the ovine model. The steady state concentrations for 17-OHPREG, PROG, and 16-OHPROG were very similar between the two models. These steady state concentrations show that an increase in the activity of the CYP17 enzyme alone increases the flux through the Δ^5 pathway, ultimately increasing the synthesis of A4. This increased CYP17 activity then also decreases the flux through the branches leading to 17-OHPROG synthesis.

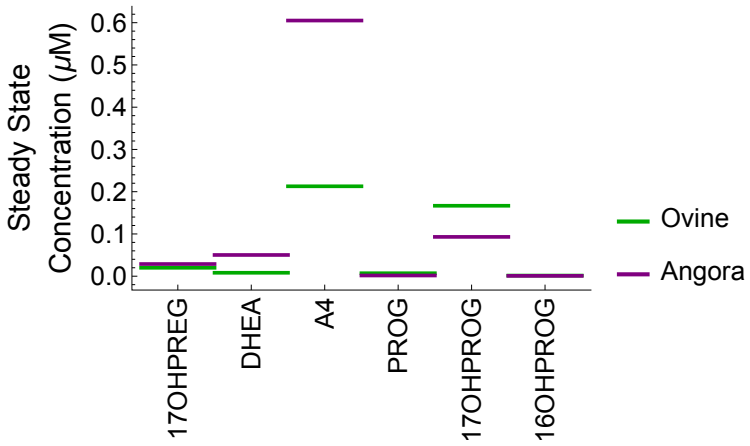


Figure 4.18: **A comparison of the steady state concentrations of 17-OHPREG, DHEA, A4, PROG, 17-OHPROG, and 16-OHPROG in the ovine and Angora models.** The steady state concentrations were calculated with a constant supply of 0.01 μM/h PREG and a constant removal of A4. The steady state concentrations of DHEA and A4 are higher in the Angora model, while the ovine model shows has higher steady state concentrations of 17-OHPROG. Steady state concentrations of PROG, 17-OHPREG, and 16-OHPROG were nearly identical between the two models.

This can also be seen in Figure 4.19, that shows the change in the fluxes to A4 and 17-OHPROG as the CYP17 and 3βHSD enzyme ratio is increased from that of the ovine model to that of the Angora model. As the CYP17 activity relative to the 3βHSD activity increases, the flux towards A4 increases, while the flux towards 17-OHPROG decreases.

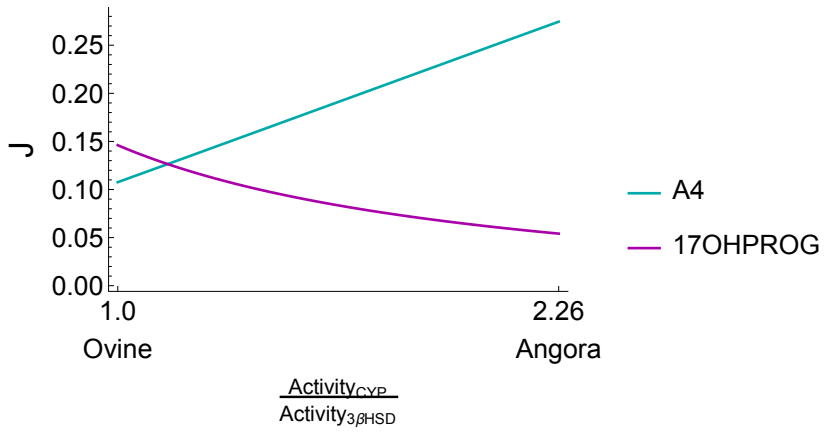


Figure 4.19: **The effect of increased CYP17 activity relative to 3βHSD activity on the fluxes towards A4 and 17-OHPROG.** The ratio of the enzyme activities are varied from that of the ovine model to that of the Angora model, leading to a decreased flux towards 17-OHPROG and an increased flux towards A4.

4.4.1 Metabolic control analysis of the ovine and Angora models

There is a clear difference in the flux through the Δ^4 and Δ^5 pathways between the ovine and Angora model (Fig. 4.19). This shift in flux is a result of the increased CYP17 activity in the Angora model. The next point to investigate is the effect that this increased CYP17 activity has on the metabolic control of the system. Metabolic control analysis (MCA), as mentioned on page 37, of the two enzymes, 3βHSD and CYP17, is shown in Figures 4.21 and 4.22. These figures are heat maps of the control coefficients of both the ovine and the Angora model shown side by side. Both CYP17 and 3βHSD catalyses multiple reactions. The effect of each of the enzymes on the flux through the pathway is considered instead of the effect of each branch of the pathway individually. Figure 4.21 shows the flux control coefficients of the enzymes on the the flux through the branches J1 to J11. In Figure 4.20 the branches of the pathway are marked according to the flux results of Figure 4.21 for easy reference. Figure 4.22 shows the concentration control coefficients of the enzymes on the concentrations of the metabolites DHEA, 17-OHPREG, 17-OHPROG, PROG, and A4. The MCA was completed with a constant supply of 0.01 μM/h PREG, and the constant removal of A4 at a rate of 0.1 h^{-1} times the metabolite concentration.

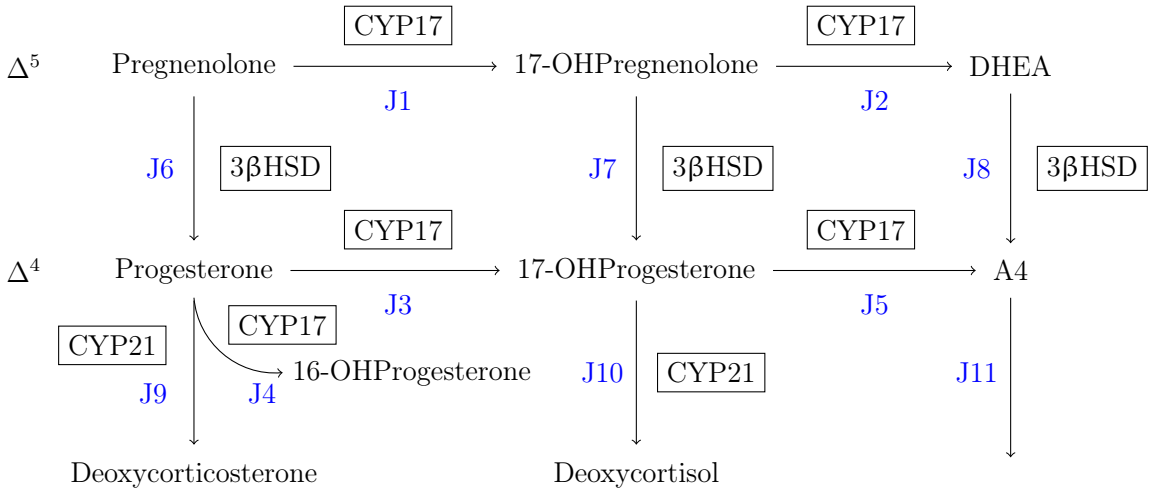


Figure 4.20: **Partial steroidogenic pathway with the flux through each of the branches numbered from J1 to J11.** Reaction J11 is the constant removal of A4 at a rate of 0.1 h^{-1} times the metabolite concentration.

Figure 4.21 shows that both enzymes have fairly similar control over the flux from PREG to 17-OHPREG (J1) in both the ovine and the Angora model. The control over the flux from 17-OHPREG to DHEA (J2) is also fairly similar between the ovine and Angora

models, with 3βHSD having little control and CYP17 having more positive control over the flux. The results are the same for the flux from DHEA to A4 (J8). There is no significant change in the control over the flux from PREG to 17-OHPREG to DHEA to A4 when the activity of the CYP17 enzyme increases from the ovine model to the Angora model. Figure 4.22 also shows that there is very little change in the control that the enzymes have on the concentrations of the metabolites 17-OHPREG, DHEA, and A4 between the two models.

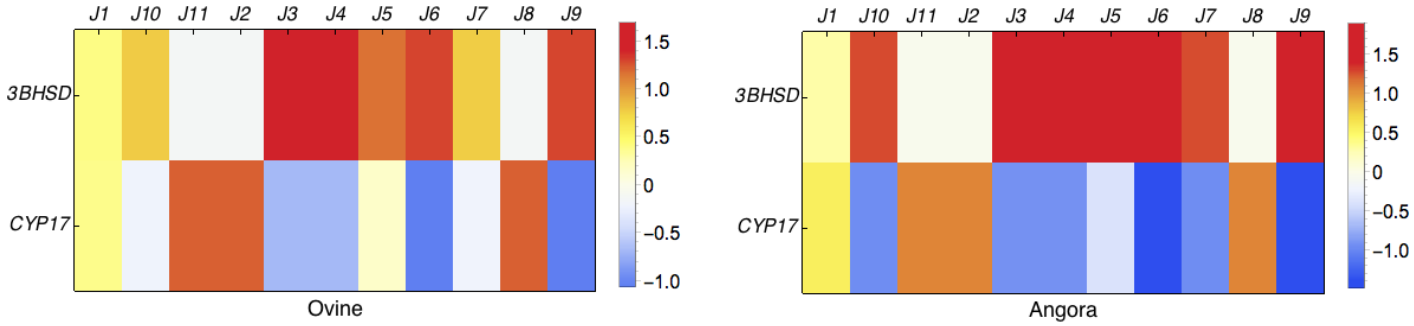


Figure 4.21: **Heat map of the flux control coefficients of the enzymes 3βHSD and CYP17 on the the flux through the branches J1 to J11** as shown in Figure 4.21.

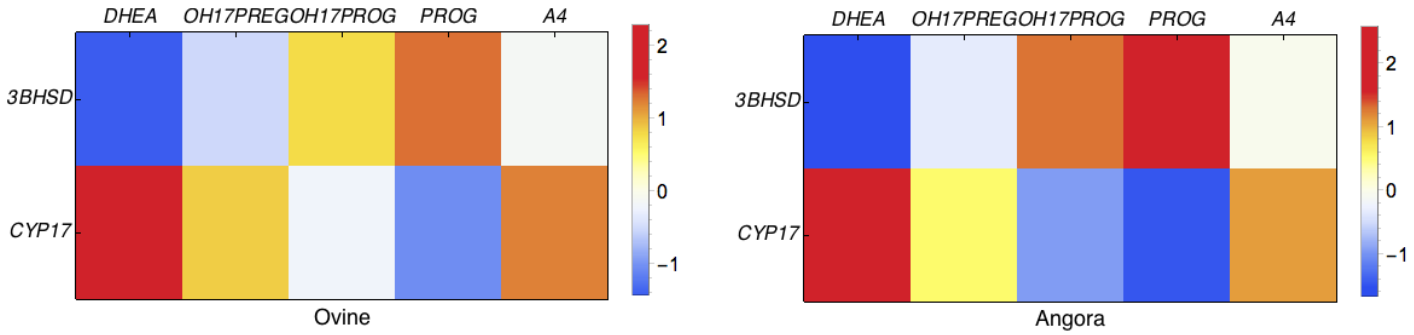


Figure 4.22: **Heat map of the concentration control coefficients of the enzymes 3βHSD and CYP17 on the concentrations of the metabolites** DHEA, 17-OHPREG, 17-OHPROG, PROG, and A4.

There is a change in the flux through the Δ^4 pathway between the ovine model and the Angora model. While 3βHSD has the same control over the flux from PROG to 17-OHPROG (J3) in both the ovine and Angora models, CYP17 has a slightly stronger negative control over this flux in the Angora model than the negative effect it has in the ovine model. The change in the control over the flux from 17-OHPROG to A4 (J5) differs the most between the two models, however the conversion of 17-OHPROG to A4 is so low in both species that this branch is often omitted during modelling [81].

The greatest change in the control over flux is for the conversion of 17-OHPREG to 17-OHPROG (J7), and the conversion of 17-OHPROG to Deoxycortisol (J10) (Fig.

4.21). The control that 3βHSD and CYP17 have over the fluxes, J7 and J10, in the ovine model increases once the activity of CYP17 increases to that of the Angora model. The increased CYP17 activity of the Angora model causes 3βHSD to have a stronger positive control over both fluxes and CYP17 to have a stronger negative control over both fluxes. Figure 4.22 show similar results, with neither enzyme having very strong control over the concentration of 17-OHPROG in the ovine model. Yet, in the Angora model, 3βHSD has a stronger positive control and CYP17 has a stronger negative control over the concentration of 17-OHPROG.

The control of both enzymes over the flux from PREG to PROG (J6) also increases, with similar results seen in Figure 4.22. 3βHSD has a stronger positive control over the flux J6 and the concentration of PROG in the Angora model than the ovine model, while CYP17 has a stronger negative control over the flux J6 and the concentration of PROG in the Angora model.

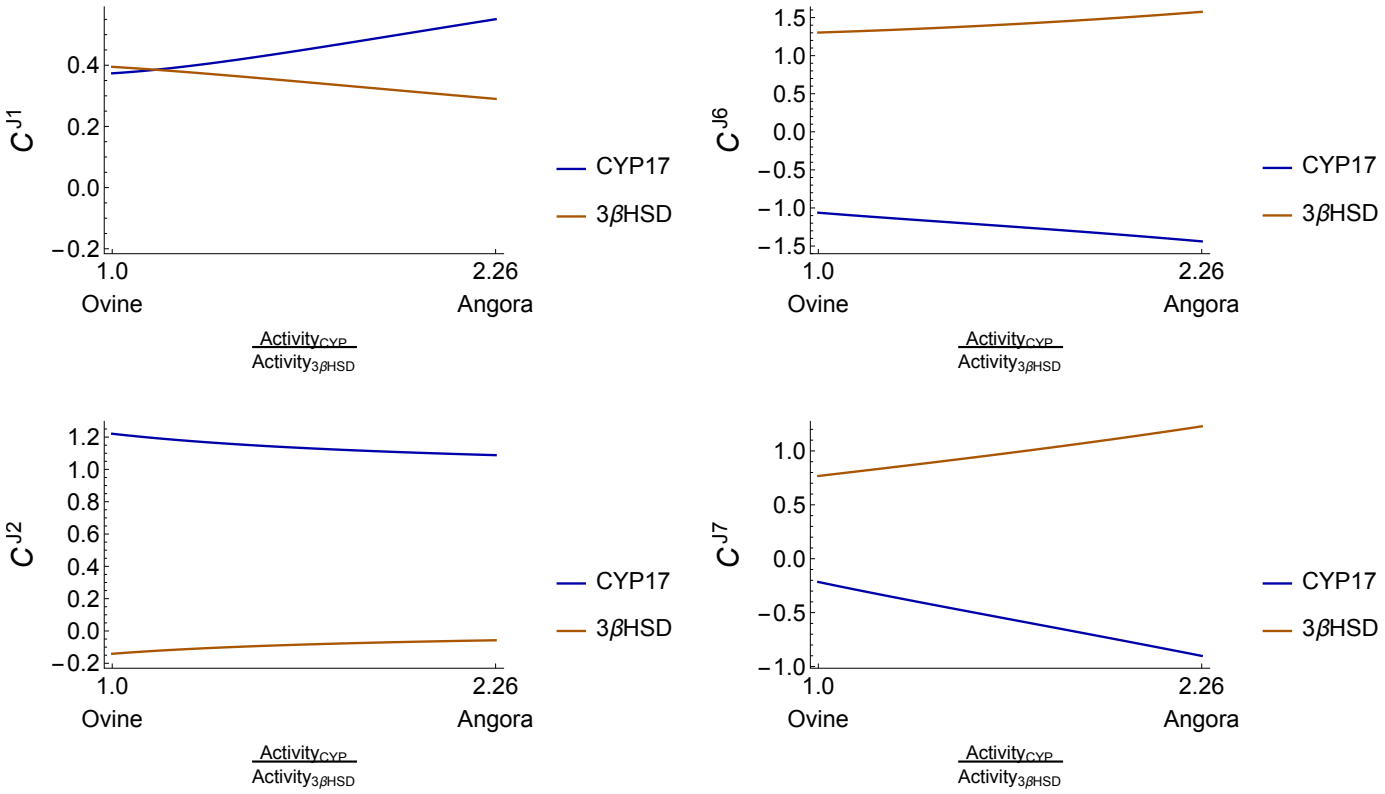


Figure 4.23: **The effect of increased CYP17 activity relative to 3βHSD activity on the control that the enzymes, CYP17 and 3βHSD, have on the fluxes J1, J2, J6 and J7.**

The changes in the flux controls between the two models are apparent in Figure 4.21, however, the quantitative changes in the system between the two models are not immediately clear. To investigate the shift in control over the fluxes of the Δ^4 and Δ^5 branches between the two models, we consider the flux pairs J1 and J6, as well as J2

and J7. These two splits in the pathway play an important role in the relative synthesis of 17-OHPROG and A4. Figure 4.23 shows the changes in the control that CYP17 and 3βHSD have on the fluxes J1, J2, J6 and J7. For both enzymes the change in the control over J6 and J7 is greater than the change in control over J1 and J2 as the enzyme activity is varied from the ration seen in the ovine model to the ratio of the Angora model.

To quantify these differences between the two models, we consider the ratio of the change in the flux control of J1 and J6, as well as that of J2 and J7, of both enzymes (see Table 4.9). The change in the control over J6 is more than the change in the control over J1 between the two models for both enzymes. Similar results are seen for the change in the control over J2 and J7, with the change in the control over J7 being higher. For both enzymes, the changes in the control over the fluxes leading to the Δ^4 branches (J6 and J7) is higher than the changes in the control over the fluxes of the Δ^5 branches (J1 and J2). The ratio of the change in flux control over J1 to the flux control over J6 (ΔC^{J1} : ΔC^{J6}) is very similar between CYP17 and 3βHSD, and so also the ΔC^{J2} : ΔC^{J7} of CYP17 and 3βHSD. There is however a difference in the flux control ratios for the split in the pathway at J1 and J6, and the split at J2 and J7. The difference between ΔC^{J1} : ΔC^{J6} and ΔC^{J2} : ΔC^{J7} is also fairly similar for both enzymes.

Table 4.9: **Flux control ratios.** These changes are the change in control over the fluxes between the ovine and the Angora models.

CYP17	3βHSD
ΔC^{J1} : ΔC^{J6} 1 : 2.13	ΔC^{J1} : ΔC^{J6} 1 : 2.60
ΔC^{J2} : ΔC^{J7} 1 : 5.19	ΔC^{J2} : ΔC^{J7} 1 : 5.51

Taken together, the shift in control between the ovine model and the Angora model is clear at both the J1 and J6, as well as the J2 and J7 split in the pathway, however, the shift in control is more prominent at the J2 and J7 split. From this one can deduce that there is not one single point in the pathway that is responsible for the shift in flux from 17-OHPROG synthesis to A4 synthesis as the CYP17 activity increases.

The following section contains the results of simulating a cold stress response with both the ovine and the Angora models.

4.4.2 Simulating the physiological stress response

The physiological response to cold stress is simulated and the differences in behaviour between the ovine and Angora model is considered. To simulate this response the concentration of PREG is increased, whereafter the steady state concentrations and changes in flux through the two models are compared.

The ratio of cortisol precursor, 17-OHPROG, and sex steroid precursor, A4 at varying concentrations of PREG is shown in Figure 4.24. With an increased concentration of PREG in the system, both models show that the synthesis of 17-OHPROG relative to A4 increases. This corresponds to the physiological response where physiological stress induces an increase in cortisol synthesis. Both models show that this response is triggered. In the ovine model, as PREG increases, the synthesis of 17-OHPROG relative to A4 increases, with more 17-OHPROG than A4 being synthesised once the concentration of PREG in the system is more than 0.2 uM. However, this response is much weaker in the Angora model. With increased PREG concentration, the synthesis of 17-OHPROG also increases, with the synthesis thereof only exceeding the synthesis of A4 once a PREG concentration of 2 uM is reached. A much stronger stress response is therefore needed for the Angora model to induce the same cortisol synthesis from increased 17-OHPROG seen in the ovine model.

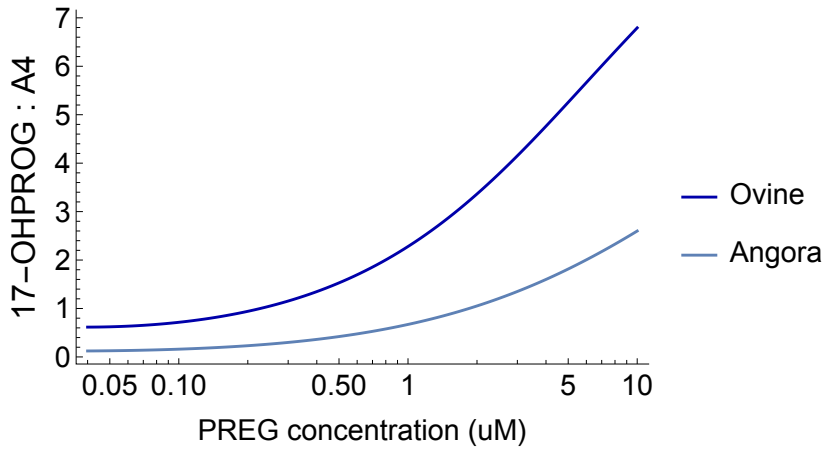


Figure 4.24: **The difference in the ratio of steady state concentrations of 17-OHPROG and A4 with increased PREG concentration between the ovine model and Angora model.**

Figure 4.25 shows the differences in the ratios of the fluxes, J1 and J6, and J2 and J7 (4.20) between the Angora and ovine models when the PREG concentration is varied. At lower PREG concentrations the flux J1 is much higher than the flux J6 in both models, however this ratio is higher in the Angora model than in the ovine model. Similar qualitative results are seen with the ratio of J2 and J7. As the PREG concentration increases, simulating the stress response in the animals, these ratios decrease, indicating that the

flux shifts from the Δ^5 pathway more towards the Δ^4 pathway in both models. Figure 4.25 shows that both species respond to the stress response by shifting the flux towards the synthesis of cortisol, this response is however much stronger in the ovine model. This also suggests that the Δ^5 pathway is favoured in the Angora species as the fluxes J1 and J2 is higher than the fluxes of the Δ^4 pathway, even at high PREG concentrations.

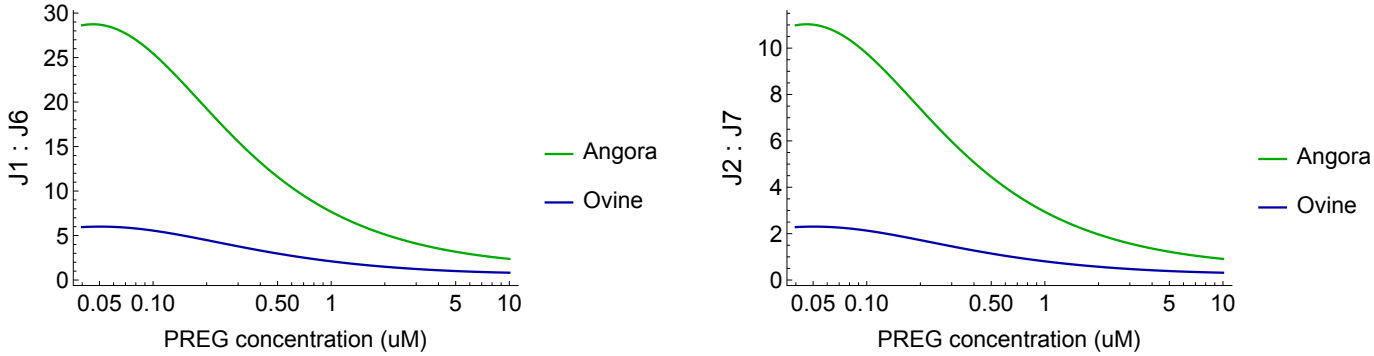


Figure 4.25: **The differences in the ratio of the fluxes J1 and J6 and the ratio of the fluxes J2 and J7 between the Angora and ovine models with increased PREG concentration.**

Taken together, with 17-OHPROG being a precursor for the synthesis of cortisol, the results of the steady state analysis and the MCA analysis support the hypothesis that the hypocortisolism seen in the Angora goats can be caused by increased CYP17 activity relative to 3βHSD activity. This is caused by the increased flux through the Δ^5 pathway, leading to increased A4 synthesis and reduced 17-OHPROG synthesis, causing a decrease in the synthesis of cortisol. As shown previously with the changes in the flux control ratios, both points in the pathway, where there is a split towards either the Δ^4 or Δ^5 pathways, contribute to the hypocortisolism in Angora goats. The point where PREG is either converted to PROG or 17-OHPREG is a point where the control of both the CYP17 and 3βHSD enzymes are increased in the Angora model. Similar results are seen at the point where 17-OHPREG is either converted to DHEA or 17-OHPROG, however, the change in the control at this point is even larger between the ovine and Angora models. The change in activity has little effect on the control over the Δ^5 branches, but has a great effect on the control over the branches leading to cortisol synthesis. This altered enzyme activity in the Angora model, reduces the flux towards cortisol precursors, and an increased flux towards the sex steroid precursors. The Angora model does respond to the simulated cold stress, however, this response is much lower than the response seen in the ovine model, maintaining a high flux through the Δ^5 branches of the pathway.

To consider the impact that these differences have on physiological level: the hypocortisolism in Angora goats are not necessarily caused by an increase in the synthesis of sex steroids, resulting in decreased cortisol synthesis, but rather a poor response to

physiological stress leading to reduced cortisol synthesis which then leads to increased sex steroid synthesis.

4.4.3 Treatment strategies

The levels of cortisol in the system regulates the blood glucose level of an animal. The current treatment administered to the young goats, in an attempt to boost their blood glucose levels, is either glucose enriched food or an intravenous dose of glucose [123, 124]. In addition to these treatments, the effect of cortisol treatments have also been investigated. It was however found that constant hypodermic treatment of young Angora goats with cortisol resulted in reduced mohair production in comparison with a control group [43]. The administering of glucose is therefore the only preventative method that can be used against hypocortisolism related deaths in Angora goats. As far as reducing the prevalence of hypocortisolism in the Angora goat species goes, the current approach is selective breeding trials with the aim of breeding a more hardy population without the loss of quality mohair production [107].

Conclusion

The ability to survive physiological stress is dependent on the levels of cortisol in the system. The low cortisol and high sex steroid precursor production seen in the South African Angora goat is most likely caused by decades of inbreeding to improve mohair quality. The development of this trade-off trait between the steroids along with improved mohair quality through the years would suggest that a high level of sex steroid production is needed to ensure the production of high quality mohair. This phenomenon is not seen in any Angora goat flocks in other parts of the world. The quality of mohair produced in countries such as Turkey, India, Albania, Russia, and the USA are inferior to South African mohair and there is no evidence of hypocortisolism among those Angora goat flocks [12, 26]. It is possible that the hypocortisolism is not just an unfortunate side effect of inbreeding, but a necessary trait for the production of superior quality mohair.

In this chapter two models were developed and validated: one for a control species with healthy steroidogenic function and one for the the Angora goat suffering from altered adrenal function. MCA showed that the increased CYP17 activity in the Angora goat changes the control that both CYP17 and 3βHSD have on the steroidogenic flux through the system. It was also found that the shift in steroidogenic flux from 17-OHPROG to A4 as enzyme activity is varied from the ovine model ratio to the Angora model ratio cannot be pin pointed to only one branch in the pathway. The point of conversion

of PREG to either PROG or 17-OHPREG and even more so the point of conversion of 17-OHPREG to either DHEA or 17-OHPROG are splits in the pathway where the steroid flux is diverted through either the Δ^4 or Δ^5 branches of the pathway. The response to cold stress by both models were then simulated and again it could be seen that these two splits in the pathway behave similarly when diverting steroid flux. A delayed response to PREG supply is seen at both of these points in the pathway. From this one can conclude that the Δ^5 pathway is the preferred pathway in the Angora goat, even with a simulated stress response, and that both the PREG and 17-OHPREG split points in the pathway contribute to the hypocortisolism phenomenon seen in Angora goats.

Chapter 5

Modelling the 5α and 5β reduction of 11-Ketotestosterone in the liver

In recent years it has been shown that 11-Ketotestosterone (11KT) is a potent androgen playing a key role in androgen dependent diseases, such as polycystic ovary syndrome (PCOS), castration resistant prostate cancer (CRPC), and congenital adrenal hyperplasia (CAH) [87, 108]. Having gone largely unnoticed, 11KT might be a more prominent androgen in women than previously thought. Unlike testosterone (T), 11KT does not decrease with age and has been shown to be prevalent androgen in PCOS [77, 87]. With the importance of 11-oxygenated androgens only recently being discovered, further study thereof is necessary.

11KT and T can either be 5α or 5β reduced. 5α reduction can be catalysed by SRD5A1 and SRD5A2, while 5β reduction is catalysed by AKR1D1 [11, 70]. The 5α reduction products or androgens are synonymous with increased androgenic activity, while the 5β reduction products are not known to show androgenic activity [112]. As such, studying the ratio of relative conversion of 11KT and T to either their 5α or 5β reduced products is of importance as the one product shows androgenic activity and the other does not.

The enzymes, AKR1D1, SRD5A1, and SRD5A2 are expressed in the liver [113–115] where they compete for the substrates, T, A4, 11KT, and 11-Ketoandrostenedione (11KA4). This chapter shows the development of a model to simulate the interaction between these three enzymes in the liver, ultimately showing whether the 5α or 5β reduction of substrates are the predominant process. The activity of AKR1D1, SRD5A1, and SRD5A2 towards the classic androgens, A4 and T, and the 11-oxygenated androgens, 11KA4 and 11KT, are characterised.

This chapter contains the model construction and parameterisation with kinetic data of transfected cells. The identifiability of the parameter values are tested and the model is validated with data from an independent transfection. The relative physiological

expression levels of these enzymes are included in the model, whereafter the model is studied under physiological steady state conditions. The model shows that the 5β reduction of 11KT is the preferred process in the liver.

5.1 Experimental data

The experimental data used for the construction of this model were collected by Lise Barnard at the Stellenbosch University. As the main focus of this chapter is the construction and further analysis of the 5α and 5β reduction model in the liver, the in depth details regarding the experimental procedure is not discussed. The detailed experimental procedure can be found in her PhD thesis and the unpublished manuscript attached in Appendix B. Experimental data were collected with the use of two sets of transfected cells. For each of the enzymes, AKR1D1, SRD5A1, and SRD5A2, two sets of experimental data were collected, one set from each transfection. The datasets collected from the two transfections differ in enzyme concentration and time range (12 and 24 hours). The data collected over 12 hours were used for model parameterisation while the data collected over a period of 24 hours were used for model validation.

5.2 Model parameterisation

The initial conditions of the experimental datasets used for model parameterisation are as follows: T (0.1 μM and 1.0 μM), 11KT (0.1 μM and 1.0 μM), A4 (0.1 μM, 1.0 μM, and 10.0 μM), and 11KA4 (0.1 μM, 1.0 μM, and 10.0 μM), for each of the three enzymes, AKR1D1, SRD5A1, and SRD5A2. The reactions, as catalysed by either one of the enzymes, are numbered from v1 to v12 in Figure 5.1. The enzyme AKR1D1 catalyses 5β reduction reactions v3, v6, v9, and v12. SRD5A1 catalyses 5α reduction reactions v1, v4, v7, and v10, while SRD5A2 catalyses the 5α reduction reactions v2, v5, v8, and v11. The 5α reduction products of the substrates A4, T, 11KA, and 11KT are 5αDione, 5αDHT, 11K5αDione, and 11K5αDHT respectively, while the 5β reduction products are 5βDione, 5βDHT, 11K5βDione, and 11K5βDHT respectively. The corresponding rate equations for these reactions, used during model parameterisation, are shown as equations 5.1 to 5.4. These rate equations depict the irreversible Michaelis-Menten kinetics of the enzymes, while the ordinary differential equations 5.5a to 5.5l depict the change in metabolite concentration over time. This ODE based model is fitted to the experimental data with nonlinear regression to determine the solutions for the parameters $Vmax_1$ to $Vmax_{12}$, and Km_1 to Km_{12}. Fitting of the model to the data is done by minimising the objective function 5.6, which is the weighted SSR value between the model and the data. The weighted nonlinear least squares fitting method was used. All experiments were

performed in triplicate, thus the SSR is weighted to the variance in the experimental data (σ_i^2). The functions `objectiveFunction` and `minimizeObjectiveFunction` of the novel Mathematica add-on package `IdentifiabilityAnalysis` was used to determine the parameters values and their respective confidence intervals. The solutions for the parameter values are shown in Table 5.1. The model fit to the experimental data are shown in Figures 5.2 to 5.4.

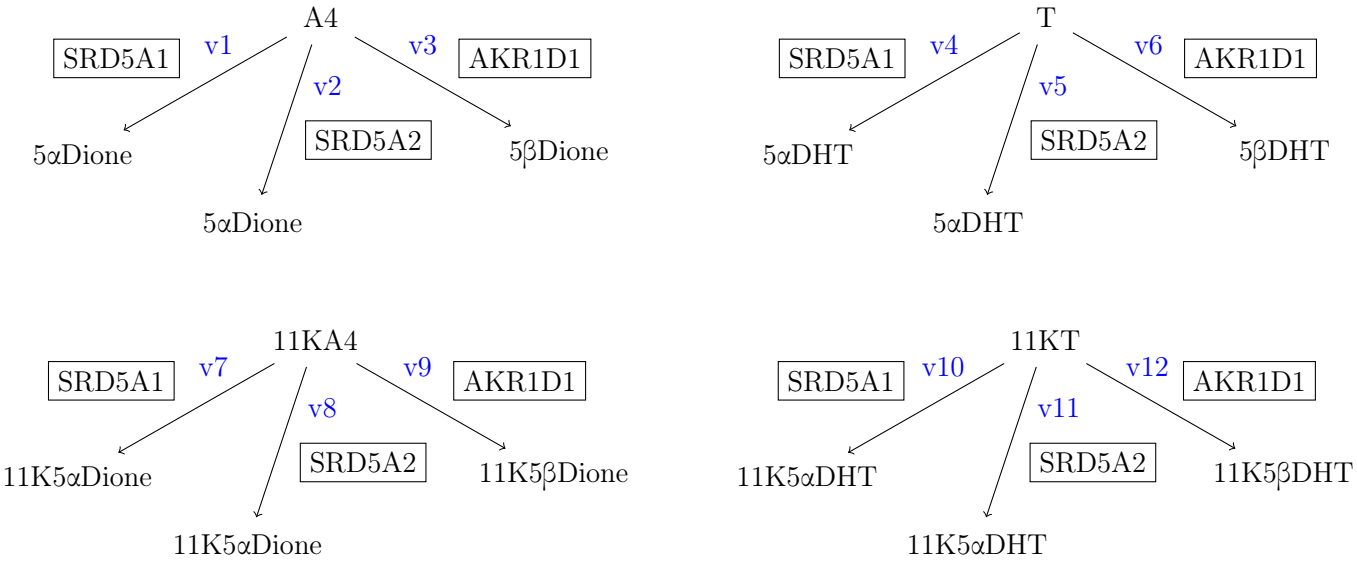


Figure 5.1: **Reactions v1 to v12, as catalysed by either AKR1D1, SRD5A1 or SRD5A2 with A4, T, 11KA4 or 11KT as substrate.**

5.2.1 Rate equations

$$v_i = \frac{Vmax_i \cdot [A4]}{Km_i + [A4]} \qquad i = 1, 2, 3 \tag{5.1}$$

$$v_j = \frac{Vmax_j \cdot [T]}{Km_j + [T]} \qquad j = 4, 5, 6 \tag{5.2}$$

$$v_k = \frac{Vmax_k \cdot [11KA4]}{Km_k + [11KA4]} \qquad k = 7, 8, 9 \tag{5.3}$$

$$v_l = \frac{Vmax_l \cdot [11KT]}{Km_l + [11KT]} \qquad l = 10, 11, 12 \tag{5.4}$$

5.2.2 Ordinary differential equations

$$\frac{d}{dt}[A4] = -v_1 - v_2 - v_3 \tag{5.5a}$$

$$\frac{d}{dt}[5\alpha Dione] = v_1 + v_2 \tag{5.5b}$$

$$\frac{d}{dt}[5\beta Dione] = v_3 \tag{5.5c}$$

$$\frac{d}{dt}[T] = -v_4 - v_5 - v_6 \tag{5.5d}$$

$$\frac{d}{dt}[5\alpha DHT] = v_4 + v_5 \tag{5.5e}$$

$$\frac{d}{dt}[5\beta DHT] = v_6 \tag{5.5f}$$

$$\frac{d}{dt}[11KA4] = -v_7 - v_8 - v_9 \tag{5.5g}$$

$$\frac{d}{dt}[11K5\alpha Dione] = v_7 + v_8 \tag{5.5h}$$

$$\frac{d}{dt}[11K5\beta Dione] = v_9 \tag{5.5i}$$

$$\frac{d}{dt}[11KT] = -v_{10} - v_{11} - v_{12} \tag{5.5j}$$

$$\frac{d}{dt}[11K5\alpha DHT] = v_{10} + v_{11} \tag{5.5k}$$

$$\frac{d}{dt}[11K5\beta DHT] = v_{12} \tag{5.5l}$$

5.2.3 Objective function

$$SSR = \sum_{i=1}^{n} \frac{(data_i - model_i)^2}{\sigma_i^2} \tag{5.6}$$

where n is the number of experimental datasets, $data_i$ is the experimental data, $model_i$ is the model fit to $data_i$ and σ_i^2 is the variance in the experimental data. During nonlinear regression the difference between $data_i$ and $model_i$, which is a function of the unknown parameters, is minimised.

Table 5.1: **Parameter values** with their respective 95% confidence intervals shown in brackets.

Parameter	μM/h	Parameter	μM
$V_{max}\ ^{A4}_{AKR1D1}$	0.029 (0.027 - 0.031)	$K_m\ ^{A4}_{AKR1D1}$	0.068 (0.052 - 0.086)
$V_{max}\ ^{11KA4}_{AKR1D1}$	0.289 (0.211 - 0.374)	$K_m\ ^{11KA4}_{AKR1D1}$	0.772 (0.454 - 1.161)
$V_{max}\ ^{T}_{AKR1D1}$	0.023 (0.015 - 0.032)	$K_m\ ^{T}_{AKR1D1}$	0.097 (0.048 - 0.168)
$V_{max}\ ^{11KT}_{AKR1D1}$	0.374 (0.279 - 0.566)	$K_m\ ^{11KT}_{AKR1D1}$	4.09 (2.98 - 6.34)
$V_{max}\ ^{A4}_{SRD5A1}$	2.57 (1.22 - 5.13)	$K_m\ ^{A4}_{SRD5A1}$	7.77 (3.28 - 15.5)
$V_{max}\ ^{11KA4}_{SRD5A1}$	1.208 (0.913 - 1.741)	$K_m\ ^{11KA4}_{SRD5A1}$	15.58 (11.11 - 23.6)
$V_{max}\ ^{T}_{SRD5A1}$	0.067 (0.047 - 0.102)	$K_m\ ^{T}_{SRD5A1}$	0.547 (0.329 - 0.985)
$V_{max}\ ^{11KT}_{SRD5A1}$	0.007 (0.004 - 0.014)	$K_m\ ^{11KT}_{SRD5A1}$	0.635 (0.232 - 1.27)
$V_{max}\ ^{A4}_{SRD5A2}$	0.254 (0.241 - 0.267)	$K_m\ ^{A4}_{SRD5A2}$	0.271 (0.234 - 0.309)
$V_{max}\ ^{11KA4}_{SRD5A2}$	0.249 (0.183 - 0.354)	$K_m\ ^{11KA4}_{SRD5A2}$	0.437 (0.252 - 0.737)
$V_{max}\ ^{T}_{SRD5A2}$	0.218 (0.195 - 0.249)	$K_m\ ^{T}_{SRD5A2}$	0.205 (0.168 - 0.254)
$V_{max}\ ^{11KT}_{SRD5A2}$	0.710 (0.455 - 1.66)	$K_m\ ^{11KT}_{SRD5A2}$	4.28 (2.59 - 10.6)

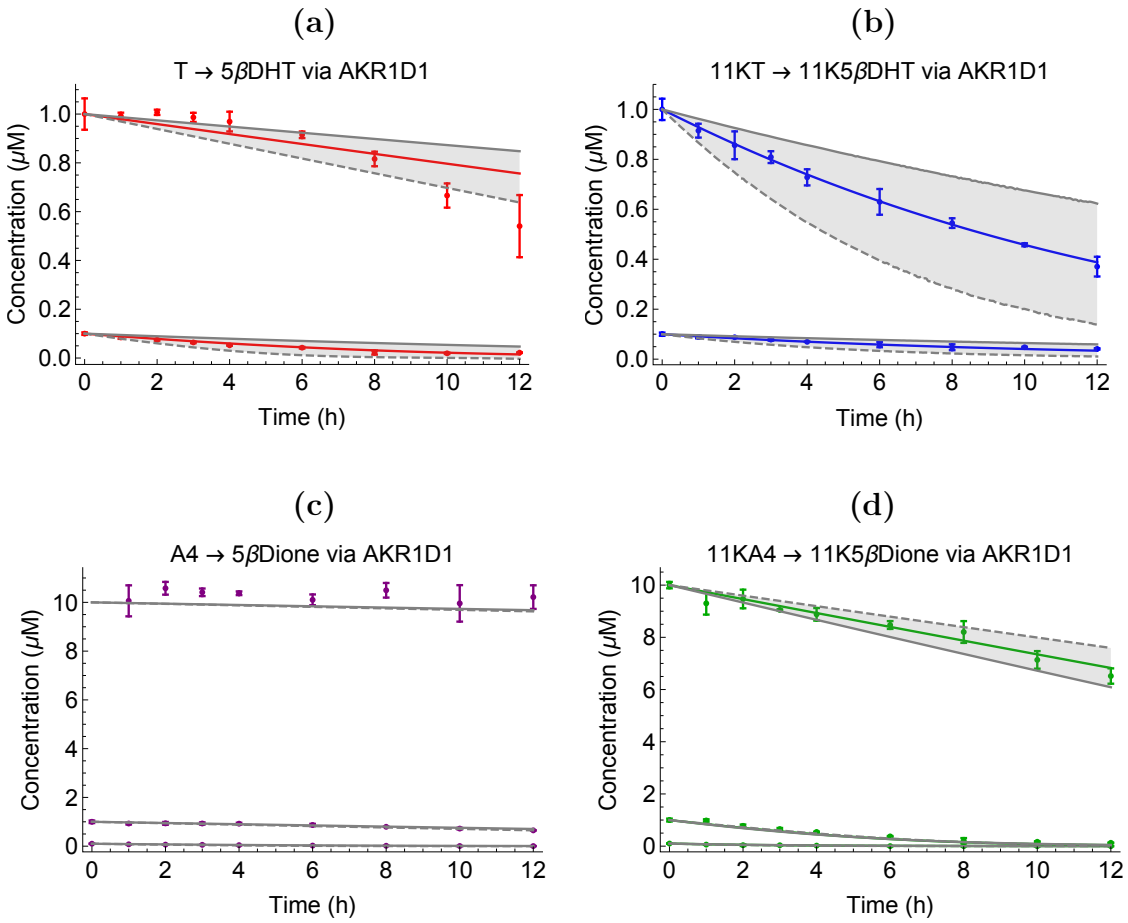


Figure 5.2: **Conversions of T to 5βDHT (a), 11KT to 11K5βDHT (b), A4 to 5βDione (c), and 11KA4 to 11K5βDione (d) by AKR1D1.** Data points are shown with their respective standard error bars. The model fits are shown as the solid lines with the 95% confidence intervals of the fits indicated by the grey areas.

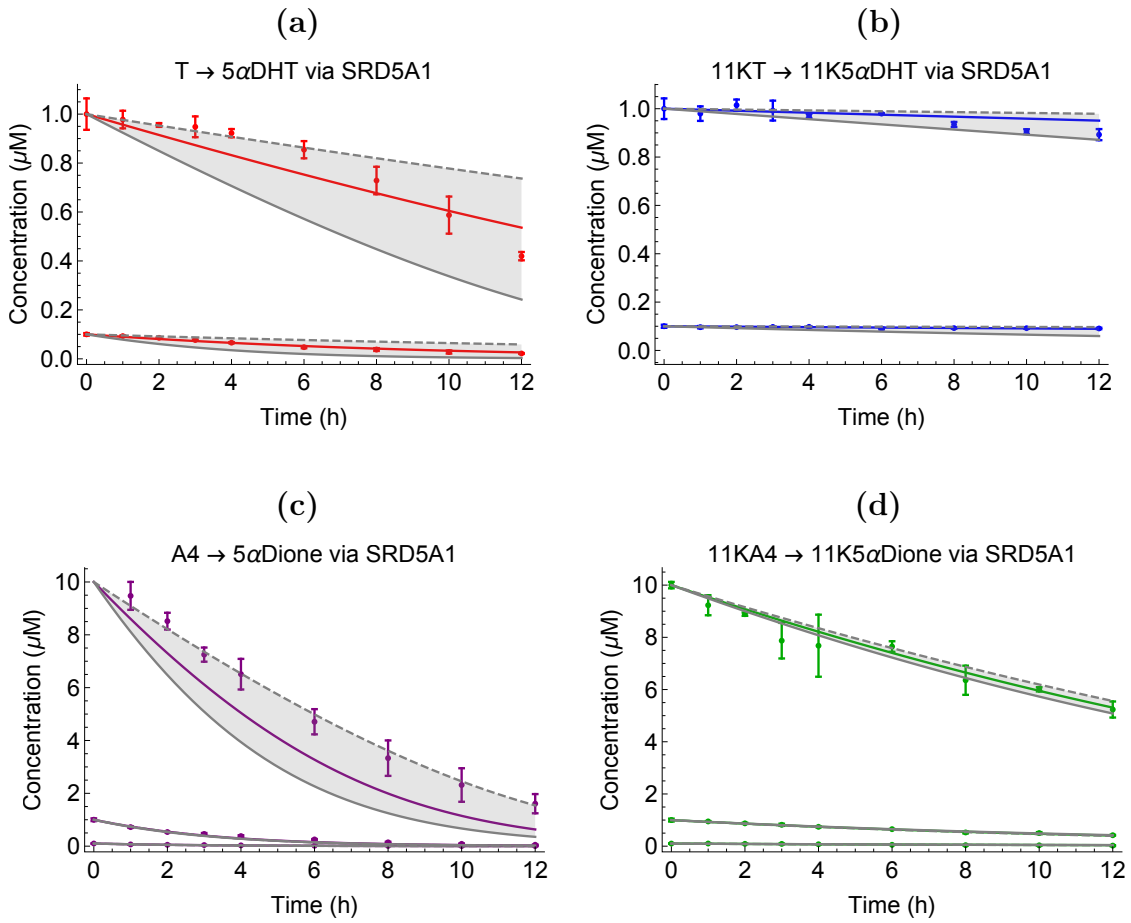


Figure 5.3: **Conversions of T to 5αDHT (a), 11KT to 11K5αDHT (b), A4 to 5αDione (c), and 11KA4 to 11K5αDione (d) by SRD5A1.** Data points are shown with their respective standard error bars. The model fits are shown as the solid lines with the 95% confidence intervals of the fits indicated by the grey areas.

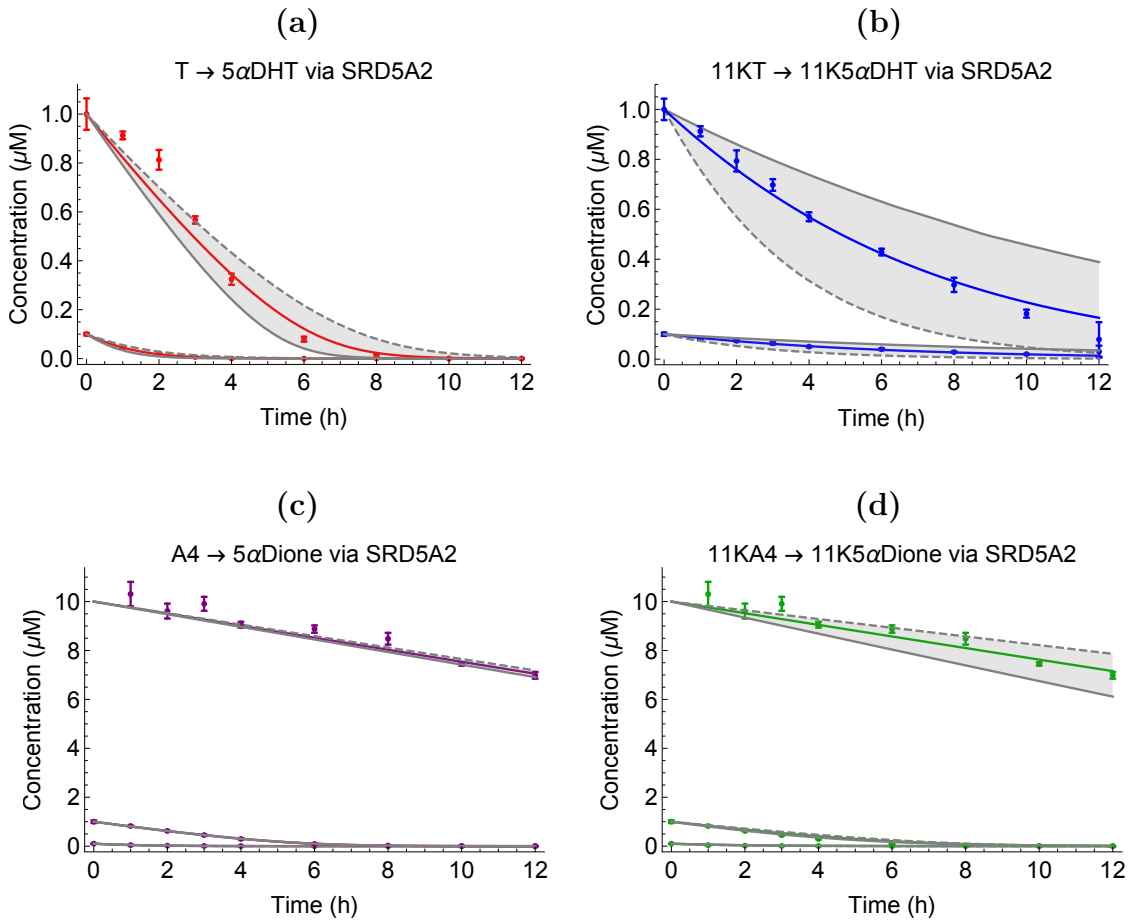


Figure 5.4: **Conversions of T to 5αDHT (a), 11KT to 11K5αDHT (b), A4 to 5αDione (c), and 11KA4 to 11K5αDione (d) by SRD5A2.** Data points are shown with their respective standard error bars. The model fits are shown as the solid lines with the 95% confidence intervals of the fits indicated by the grey areas.

5.3 Identifiability analysis

The identifiability of all the parameter values were tested. The likelihood ratio test was used to create profile likelihood plots and determine the 95% confidence intervals as well as whether the parameter values are identifiable [51, 89]. The profile likelihood values are indicated by the solid black lines in the profile likelihood plots. The 95% Chi-squared distribution value is indicated by the solid red threshold line on the plots. The optimal parameter value solution, as determined during model fitting, is indicated by the dashed blue line. The parameter value is enclosed within the 95% confidence interval, as shown by the solid blue lines. These confidence intervals represent the value where the profile likelihood of the parameter intercepts the 95% Chi-squared threshold. All of the estimated values were identifiable except for the V_{max} and K_m values for the conversion of 11KT to 11K5αDHT, and A4 to 5αDione by SRD5A1. These values were practically non-identifiable, with a 95% confidence interval that is lower bounded, but not upper bounded. For these non-identifiable parameters a value of two times the parameter value was used as the upper limit of the confidence interval. The profile likelihood plots of the parameters are included at the end of this document in Appendix B. These profile likelihood plots were created with the functions `profileLikelihood` and `likelihoodRatioPlot` of the Mathematica add-on package `IdentifiabilityAnalysis`, see chapter 3.

5.4 Model validation

The parameter value estimations were validated by using the model to predict the outcome of the experiments completed with another transfection. The initial conditions of the experimental datasets used for model validation are as follows: T (0.1 µM, 1.0 µM, and 10.0 µM) and 11KT (0.1 µM, 1.0 µM, and 10.0 µM), for each of the three enzymes, AKR1D1, SRD5A1, and SRD5A2. A variable multiplier was fitted for to account for the transfection efficiency of each of the enzymes. The variable multiplier values fitted for for each of the enzymes are 1.55, 2.34, and 0.667 for AKR1D1, SRD5A1, and SRD5A2 respectively. Only one multiplier value was fitted for per enzyme for both the T and 11KT substrate conversions during the prediction of the experimental time course data. These multipliers correct for the relative levels of enzyme expression and is included as such in the rate equations, while all parameter values were kept fixed. The model predictions of the validation data are shown in Figure 5.5. These variable multipliers only adjust for the transfection efficiency and are used, along with the determined $Vmax$ and Km values of each enzyme to predict the validation datasets.

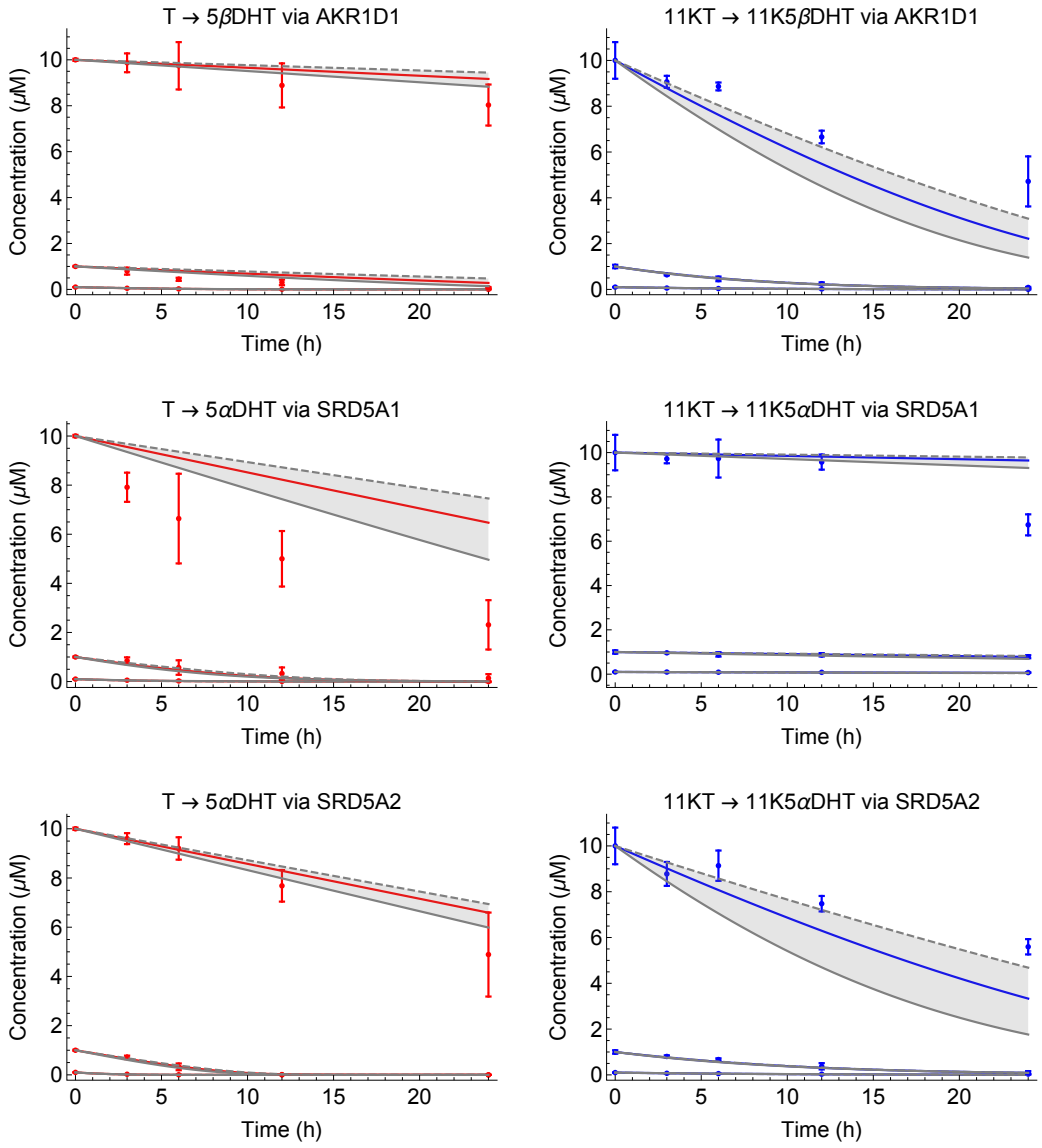


Figure 5.5: **Validation of newly determined AKR1D1, SRD5A1, and SRD5A2 parameter values.** Model predictions (solid lines) of the experimental data (points) are shown.

5.5 Modelling of AKR1D1, SRD5A1 and SRD5A2 kinetics with 1:1:1 enzme expression

The parameter values of AKR1D1, SRD5A1, and SRD5A2, as shown in Table 5.1, were used to construct a model which describes equal expression of the three enzymes (in a ratio of 1:1:1). This undetermined model can be used to simulate the conversion of any initial concentrations of T, A4, 11KT, and/or 11KA4 at any expression levels of the enzymes AKR1D1, SRD5A1, and SRD5A2. In addition to the parameters listed in Table 5.1, three parameters were added to the model. The three parameters (`AKR1D1exp`,

`SRD5A1exp`, and `SRD5A2exp`), which represent the expression levels of the three enzymes, allows one to change the enzyme expression levels to any desired values.

Experimental results of mixing of the transfected enzymes in a 1:1 volume ratio were predicted with the model as validation thereof. It is assumed that the transfection efficiencies of the transfected enzymes are similar. With the same initial concentrations under which the experimental data were collected, model predictions of the experimental results were simulated. These model predictions and experimental results can be seen in Figures 5.6 and 5.7. AKR1D1 was mixed in a ratio of 1:1 with either SRD5A1 or SRD5A2. The conversion of 1 μM T, 11KT, A4 or 11KA4 to their respective products were measured after 24 hours.

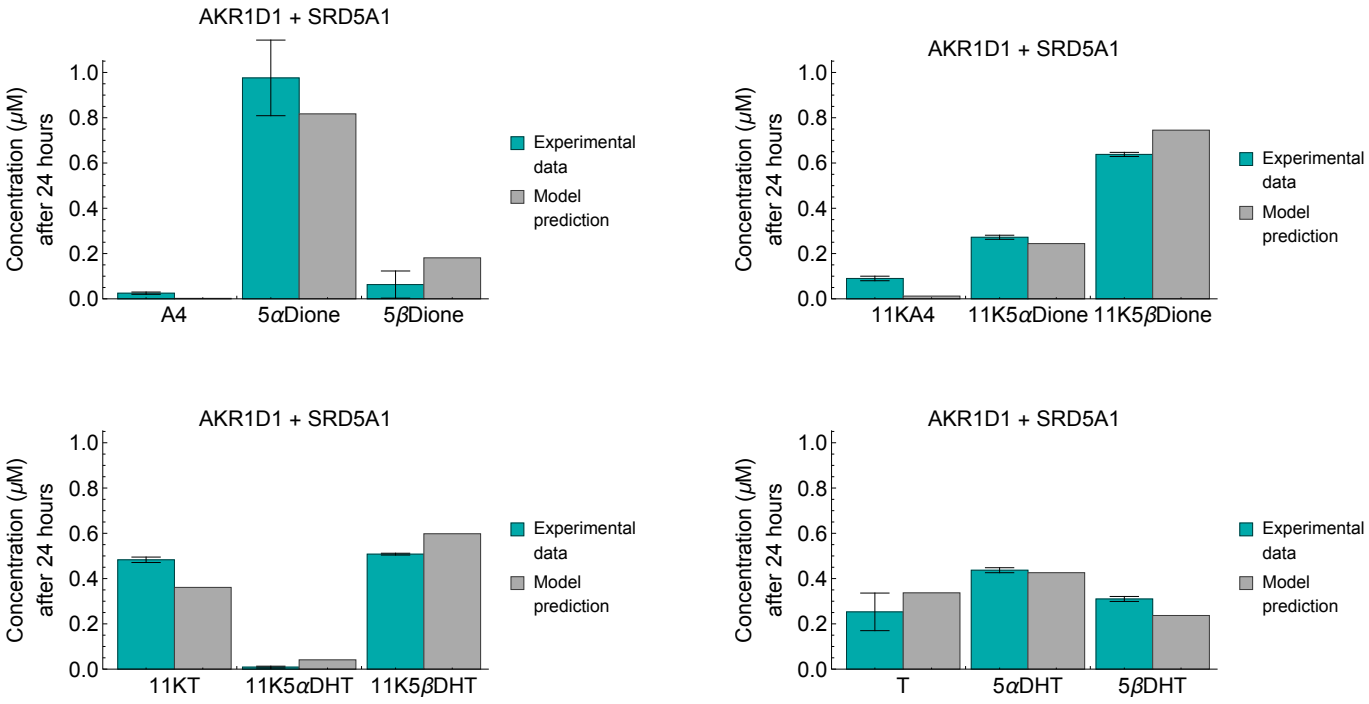


Figure 5.6: **Model prediction of the experimental results of 1:1 mixing of AKR1D1 with SRD5A1.** The conversion of 1 μM T, 11KT, A4 or 11KA4 were measured after 24 hours.

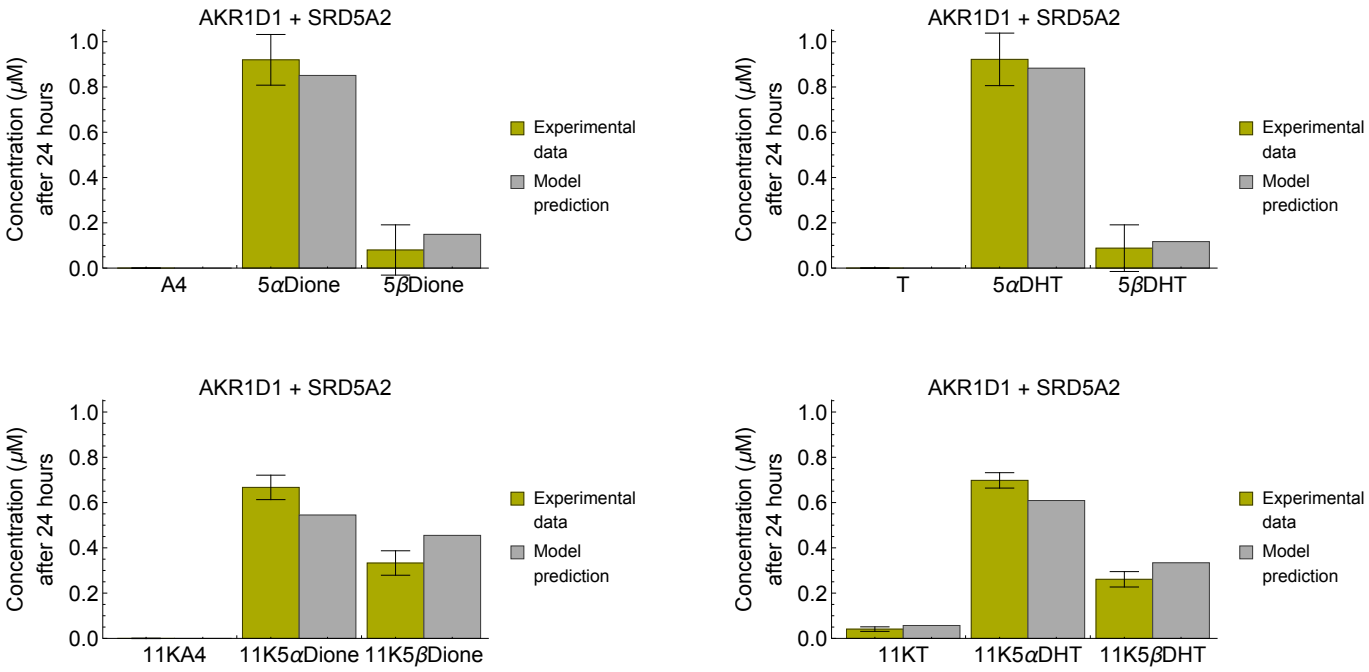


Figure 5.7: **Model prediction of the experimental results of 1:1 mixing of AKR1D1 with SRD5A2.** The conversion of 1 μM T, 11KT, A4 or 11KA4 were measured after 24 hours.

Figure 5.8 shows the change in enzyme activities with variation in T and 11KT concentration. The activity of SRD5A2 is the highest of the three enzymes at all concentrations of 11KT. SRD5A1 has the highest affinity for 11KT, but has a very low maximal rate. AKR1D1 and SRD5A2 have similar affinities for 11KT, however SRD5A2 has a higher maximal rate, resulting in SRD5A2 having a higher catalytic efficiency (V_{max}/K_m) towards 11KT than AKR1D1. With T as substrate, SRD5A2 again has the highest activity at all concentrations of T. At lower concentrations of T (up to 0.15 μM) AKR1D1 has higher activity than SRD5A1, however, above that threshold, the activity of SRD5A1 is higher. When considering the catalytic efficiencies of the enzymes with T as substrate, the efficiency of SRD5A2 is the highest, with a value of 1.063, followed by AKR1D1, with a value of 0.237, with the efficiency of SRD5A1 being the lowest at 0.122.

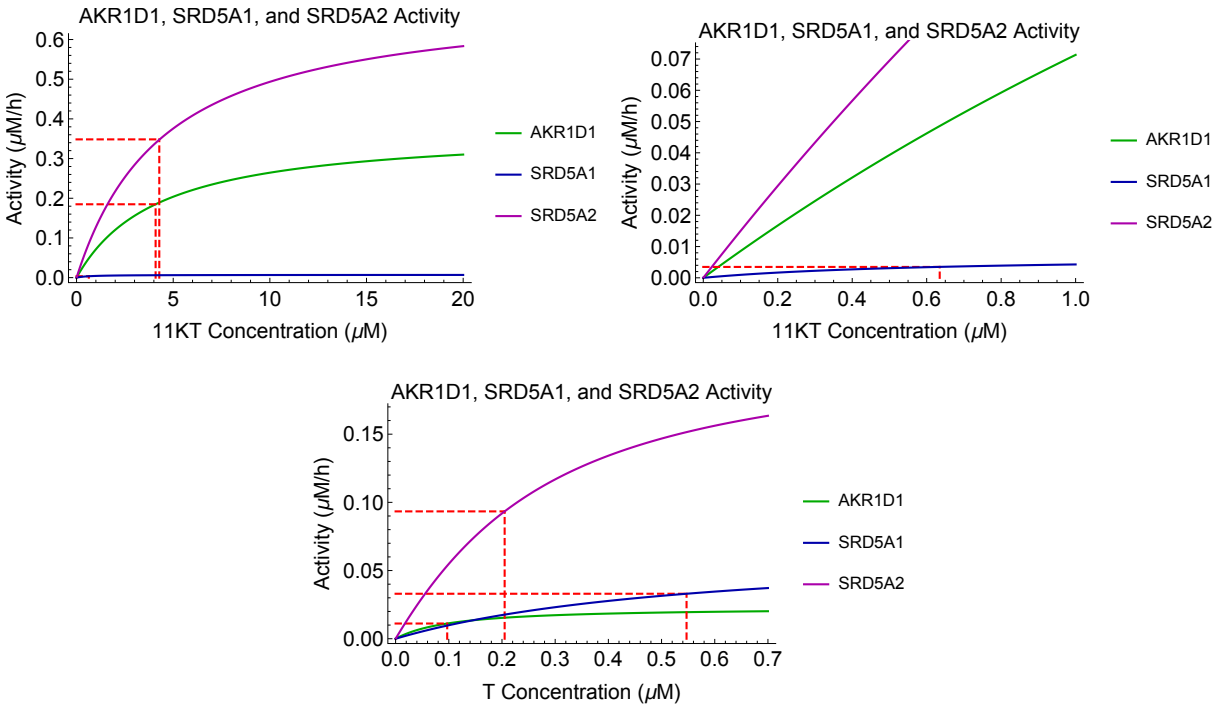


Figure 5.8: **Activities of AKR1D1, SRD5A1, and SRD5A2 with variation of substrate (11KT and T) concentration.**

5.6 Modeling of AKR1D1, SRD5A1 and SRD5A2 kinetics in the liver

To simulate the physiological expression of AKR1D1, SRD5A1, and SRD5A2 in the liver, the relative RNA expression of these enzymes were used. The absolute values of 255.7 [113], 16.3 [114] and 34.2 [115] tags per million, as recorded on the The Human Protein Atlas database were used. A limitation of using the RNA expression levels of these enzymes as an indicator for the protein expression levels is that the relative RNA levels do not always necessarily relate to the relative enzyme expression levels. At this point these were the most accurate data available. The parameters `AKR1D1exp`, `SRD5A1exp`, and `SRD5A2exp` were set to the scaled, relative values of 15.7, 1, and 2.10 for AKR1D1, SRD5A1, and SRD5A2 respectively.

Figure 5.9 shows model simulations of the conversion of 0.8 nM 11KT to 11K5αDHT and 11K5βDHT, and 0.8 nM T to 5αDHT and 5βDHT at enzyme expression levels of 1:1:1 in comparison to physiological enzyme expression levels. As experimental data of the synthesis of the 5α and 5β reduction products of the substrates 11KT, 11KA4, T, and A4 with physiological enzyme expression levels has not yet been collected, model simulations can be used to predict the relative synthesis of these products. The initial concentrations of 0.8 nM are close to physiological plasma concentrations of T and 11KT

[52]. At enzyme expression levels in the ratio 1:1:1 the 5α reduction of both T and 11KT is the preferred reaction. With T as substrate, being fully converted to either 5αDHT or 5βDHT, the majority of T is converted to 5αDHT (83.4%) with only 16.6% of T being converted to 5βDHT. Although the 5α reduction product is also the predominant product formed with 11KT as substrate, more of the 5β reduction product is formed than with T as substrate, with 65.9% 11K5αDHT and 34.1% 11K5βDHT being synthesised. This is however not the case at physiological expression levels of these enzymes. As the physiological expression level of AKR1D1 is higher than that of SRD5A1 and SRD5A2, the 5β reduction of both T and 11KT are preferred. If all T substrate is fully converted to 5αDHT and 5βDHT, about 38.8% of T is converted to 5αDHT and about 61.2% T is converted to 5βDHT. With 11KT as substrate much more of the 5β reduction product, 5βDHT, is synthesised (80.0%), while less of the 5α reduction product, 5αDHT, is synthesised (20.0%). Similar qualitative results are seen for the conversion of 11KA4 to 11K5αDione and 11K5βDione, as well as the conversion of A4 to 5αDione and 5βDione.

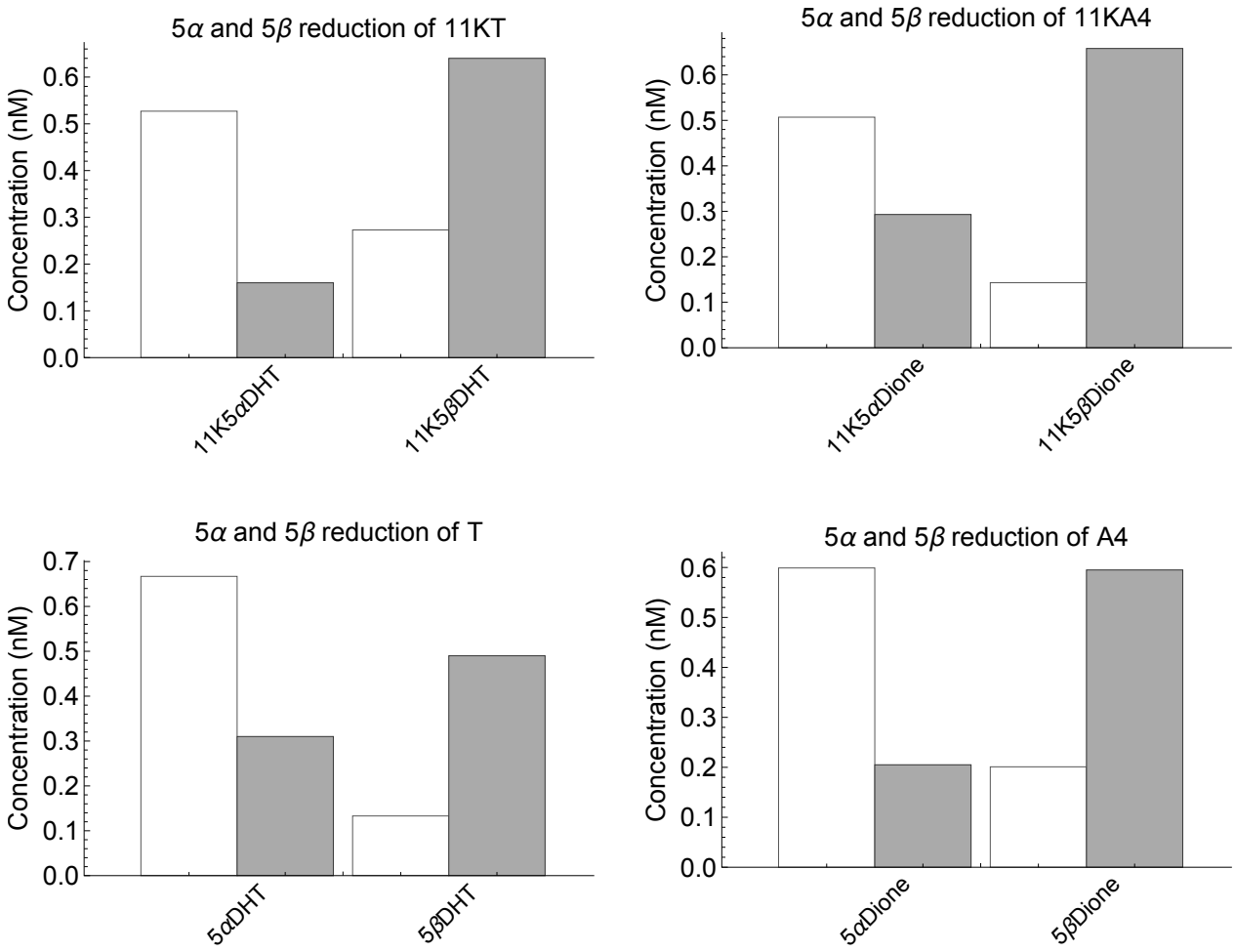


Figure 5.9: **Model simulations of the conversion of 0.8 nM 11KT, T, 11KA4 or A4** to their respective 5α and 5β reduction products at 1:1:1 enzyme expression (white) and physiological enzyme expression levels (gray).

5.6.1 Steady state analysis with T and 11KT as substrates

The dynamics of the model under the following steady state conditions were studied: a constant inflow of 0.0720 μM/h and 0.0648 μM/h 11KT into the liver for men and women respectively and a constant inflow of T into the liver of 1.419 μM/h for men and 0.0632 μM/h for women. This is calculated from the plasma levels of 234 pg/mL 11KT and 4548 pg/mL T in men and 250 pg/mL 11KT and 225 pg/mL T in women [52] and liver perfusion rates of 1.35 L/min in women and 1.5 L/min in men [60]. The degradation of steroid products or removal thereof by the system is included in the model as 0.1 min^{-1} times the steroid concentration, similar to the steroid hormone modelling techniques used by Nguyen *et al* [81]. Only the initial 5α or 5β reduction of the steroid hormones are considered and no further conversion of the 5α and 5β products.

The change in steady state concentrations of the 5α and 5β reduction products 11K5αDHT, 11K5βDHT, 5αDHT, and 5βDHT with the variation in AKR1D1 expression relative to SRD5A1 and SRD5A2 expression can be seen in Figure 5.10. Under these steady state conditions, which were chosen to resemble a physiological steady state, the 5β reduction of 11KT is the preferred reaction once AKR1D1 is expressed in the liver at a ratio of 2:1:1 or higher relative to SRD5A1 and SRD5A2, while the 5β reduction of T is only preferred once that ratio is 5:1:1 or higher.

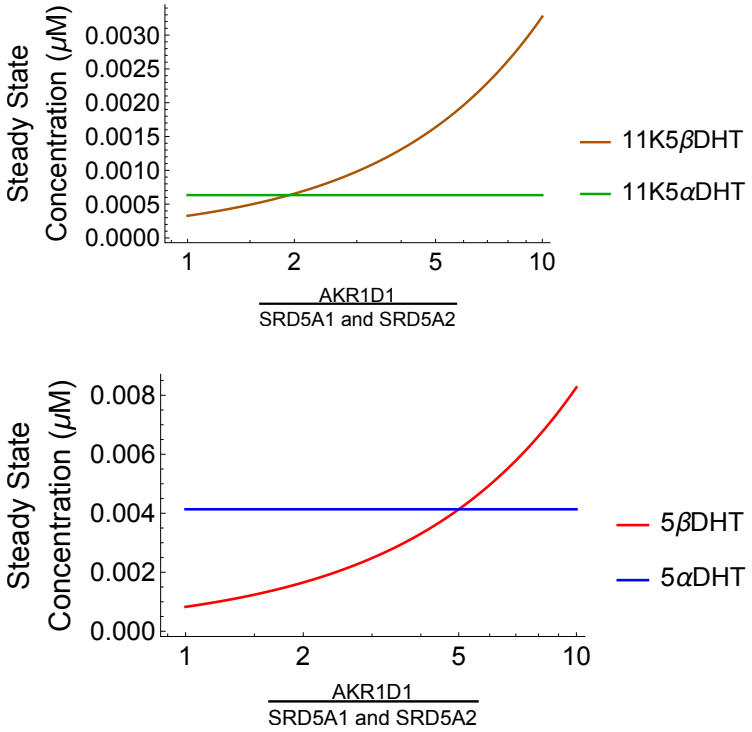


Figure 5.10: **Variation in steady state concentrations of 11K5αDHT, 11K5βDHT, 5αDHT, and 5βDHT as the relative expression of AKR1D1 to SRD5A1 and SRD5A2 changes.**

The metabolites T, 11KT, A4, and 11KA4 are pooled *in vivo* and the enzymes AKR1D1, SRD5A1, and SRD5A2 compete for the same pooled substrates. The plasma levels of T in men are higher than in women, which not only influences the synthesis of 5αDHT and 5βDHT, but also that of 11K5αDHT and 11K5βDHT. This large difference in T plasma levels has little effect on the synthesis of 5αDHT relative to 5βDHT, and the synthesis of 11K5αDHT relative to 11K5βDHT. In both men and women 5αDHT and 5βDHT are synthesised in a ratio of approximately 1:1.75, while 11K5αDHT and 11K5βDHT are synthesised in a ratio of approximately 1:4. The absolute values of the steroid hormones do however differ between men and women, but this is expected due to the much higher levels of T in men. As shown previously, the 5β reduction of 11KT is the preferred reaction with physiological enzyme expression levels. Steady state concentrations of the metabolites 11K5αDHT, 11K5βDHT, 5αDHT, and 5βDHT for men and women are shown in Figure 5.11.

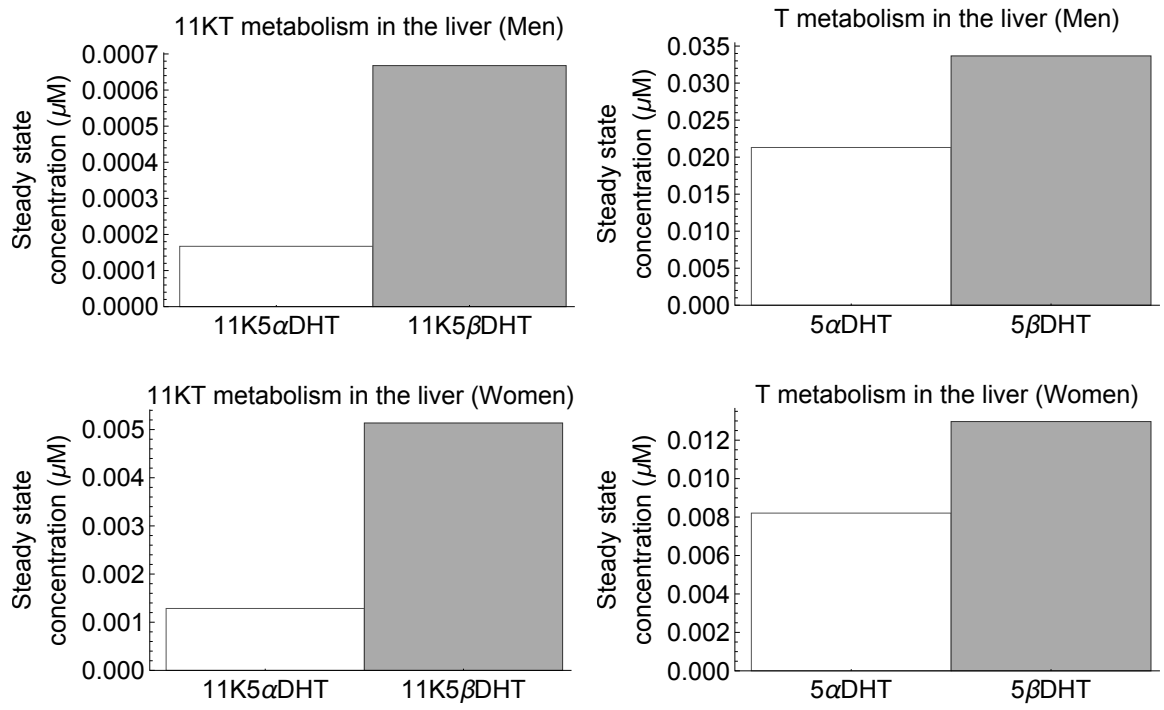


Figure 5.11: **Model simulation of the steady state concentrations of 11K5αDHT, 11K5βDHT, 5αDHT, and 5βDHT in the liver with physiological expression of AKR1D1, SRD5A1, and SRD5A2 (15.7:1:2.1).**

Conclusion

Knowing the relative conversion of androgens such as T, 11KT, A4, and 11KA4 to their respective 5α and 5β reduction products is of importance, as the 5α reduction products show strong androgenic activity, while the 5β reduction products do not. The enzymes

AKR1D1, SRD5A1, and SRD5A2 are expressed in the liver, with AKR1D1 catalysing the 5β reduction of androgens and SRD5A1 and SRD5A2 catalysing the 5α reduction of androgens. These enzymes compete for the same substrates and have different expression levels. The enzymes also have different efficacies for the androgen substrates. It is therefore useful to have a mathematical model which describe these enzyme dynamics and the relative synthesis of 5α and 5β reduction products. We created a model that can simulate these steroid conversions in the liver. The model was parameterised with experimental data form reconstituted cells, whereafter it was validated with experimental data obtained from mixing cells expressing AKR1D1 and cells expressing either SRD5A2 or SRD5A2. The model has been constructed such that the steroid conversion, given any rate of androgen inflow into the liver, at any levels of enzyme expression can be simulated. We studied the specific case of the model, where the enzymes are expressed in the physiological ratio and see that the 5β reduction of 11KT is the predominant reaction in the liver. Similarly, the 5β reduction of the other androgens are also the predominant reactions in the liver under physiological conditions of both men and women.

Chapter 6

Computational modelling of the intratumoral androgen metabolism in castration resistant prostate cancer

Prostate cancer is one of the most prevalent forms of cancer in men. Androgen deprivation therapy via chemical or physical castration is the current treatment for prostate cancer. Despite this treatment castration resistant prostate cancer (CRPC) often emerges [22, 98]. Prostate cancer cells express enzymes that convert androgen precursors from the adrenal glands to potent androgens [45].

Under healthy conditions A4 is converted to T via AKR1C3 which is in turn converted to 5αDHT via SRD5A (see Figure 6.1). T can also be converted back to A4 via 17βHSD2. Androgen depravation therapy significantly lowers the amount of T in circulation, therefore limiting the synthesis of potent 5αDHT. What is seen with the emergence of CRPC is that 5αDHT is synthesised from A4 with 5αDione as intermediate; A4 is converted to 5αDione via SRD5A and 5αDione is in turn converted to 5αDHT via AKR1C3. 17βHSD2 can also convert 5αDHT back to 5αDione. The synthesis of 5αDHT with 5αDione as intermediate is known as the alternate 5αDione pathway. This pathway is however not the only source of androgens for CRPC. It has been found that 11β-hydroxyandrostenedione (11OHA4) is a precursor for the potent androgens 11KT and 11K5αDHT, see Figure 6.2, this pathways is known as the 11OHA4 pathway. 11K5αDHT is equipotent to 5αDHT as a substrate for the enzymes in prostate cancer cells. It was also found that 11KA4 and 11K5αDione are better substrates for AKR1C3 than A4 or 5αDione [7]. The 11OHA4 pathway has not yet been studied in as much detail as the classic androgen pathway or the 5αDione pathway. The same enzymes are present in both the 5αDione and the 11OHA4 pathway and both pathways are active in CRPC.

The classic androgens and the 11-oxygenated androgens are pooled within the CRPC tumors, where the enzymes can catalyse either the reactions of the 5αDione or the 11OHA4 pathway. Due to the multi-functionality of these enzymes and the different effects of the classic and the 11-oxygenated androgens, it is of importance to determine the relative contribution of the classic and 11-oxygenated androgens to the combined androgen pool.

In an attempt to predict the contribution of either group of androgens to the androgen pool, we have constructed a model that describes the dynamics of the 5αDione biosynthesis pathway as shown in Figure 6.1 and the dynamics of the 11-oxygenated androgen biosynthesis pathway as shown in Figure 6.2 in CRPC C42B cells. The combined dynamics of the AKR1C3 and 17βHSD2 enzymes have previously been parameterised, these values will be taken from literature [7]. All data used in this chapter for model construction and validation were collected by Prof Storbeck and members of his group at Stellenbosch University, including the data published in [7]. The SRD5A data have previously been analysed by our group and the dynamics thereof have been parameterised. These unpublished results will be used as the parameter values for SRD5A.

The dynamics of the 11βHSD2 and 3αHSD enzymes are parameterised and validated in this chapter and these newly determined parameter values are used along with the other, previously determined parameter values, to construct the model. The parameter values for the kinetics of 11βHSD2 are determined from time course data of transfection experiments. No kinetic data for the reactions catalysed by 3αHSD have been collected, therefore, the experimental data from C42B prostate cancer cells have been used to fit for and validate the 3αHSD parameters. Metabolic control analysis identifies the enzymes with most control over hormone flux through the system. The effect of increased expression of the enzymes on the steady state concentrations of the metabolites are investigated. Lastly, steady state analysis of the model shows the results of enzyme inhibition and which enzymes are possible treatment targets for CRPC.

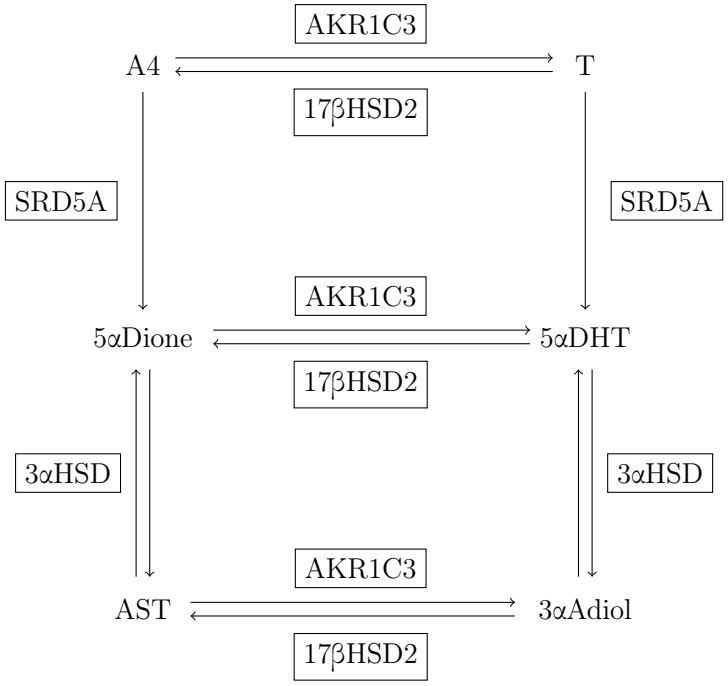


Figure 6.1: **Schematic representation of the classic androgen (A4 to 5αDHT via T) and 5αDione biosynthesis pathway (A4 to 5αDHT via 5αDione).**

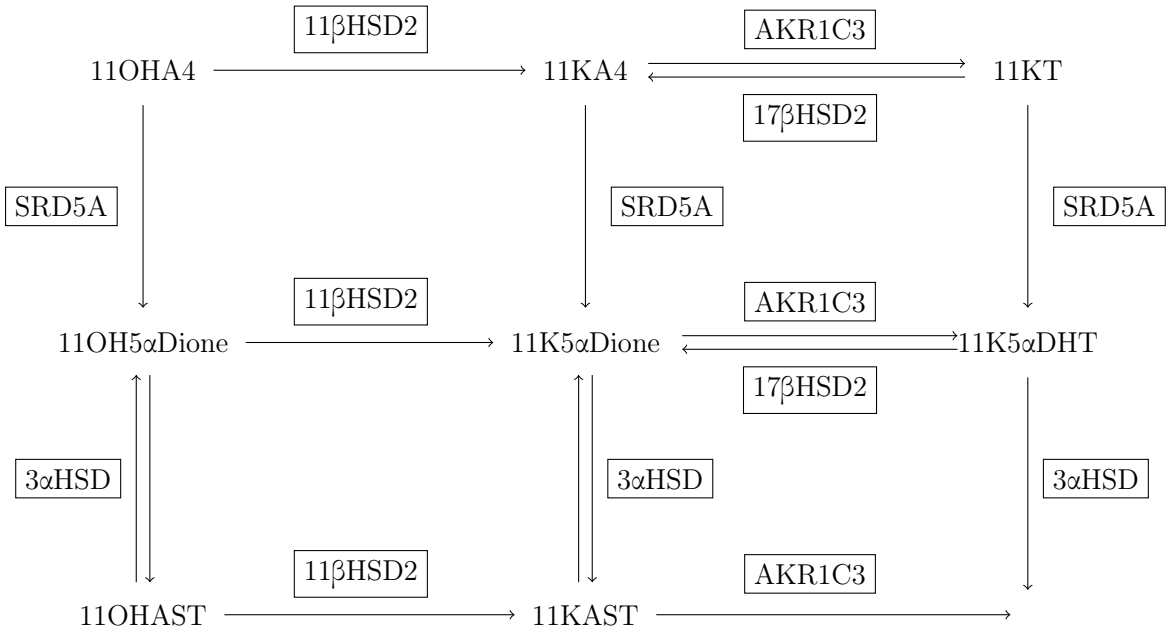


Figure 6.2: **Schematic representation of the 11-oxygenated androgen biosynthesis pathway.**

6.1 Previously determined parameter values

The parameter values for the enzymes AKR1C3, 17βHSD2 [7], and SRD5A have previously been determined (unpublished results) and were used as is in all further analyses (values shown in Tables 6.1 to 6.3).

Table 6.1: **Parameter values for AKR1C3 kinetics** with their respective 95% confidence intervals shown in brackets.

Substrate	V_{max} (μM/h)	K_m (μM)
11KA4	0.27 (0.25 - 0.30)	0.11 (0.07 - 0.13)
11K5αDione	0.55 (0.44 - 0.72)	0.23 (0.16 - 0.33)
11KAST	0.59 (0.48 - 0.78)	0.34 (0.24 - 0.50)
A4	0.02 (0.01 - 0.02)	0.06 (0.03 - 0.09)
5αDione	0.01 (0.01 - 0.02)	0.12 (0.08 - 0.20)
AST	0.05 (0.04 - 0.08)	0.27 (0.14 - 0.55)

Table 6.2: **Parameter values for 17βHSD2 kinetics** with their respective 95% confidence intervals shown in brackets.

Substrate	V_{max} (μM/h)	K_m (μM)
T	3.71 (3.41 - 4.04)	5.08 (4.56 - 5.67)
5αDHT	2.25 (2.07 - 2.44)	3.77 (3.40 - 4.17)
3αAdiol	3.99 (3.49 - 4.57)	5.86 (4.94 - 7.03)
11KT	4.35 (4.01 - 4.73)	9.11 (8.24 - 10.12)
11K5αDHT	6.40 (5.70 - 7.23)	7.72 (6.67 - 9.02)

Table 6.3: **Parameter values for SRD5A kinetics** with their respective 95% confidence intervals shown in brackets.

Substrate	V_{max} (μM/h)	K_m (μM)
A4	0.0066 (0.0058 - 0.0079)	12.1 (9.97 - 14.76)
T	0.0936 (0.0414 - 0.187)	373 (157 - 762)
11OHA4	0.00378 (0.0031 - 0.0047)	30.4 (24.1 - 40.1)
11KA4	0.0156 (0.0078 - 0.0318)	90.7 (41.2 - 187)
11KT	0.021 (0.003 - 0.042)	918 (118 - 1847)

6.2 Newly determined parameter values

6.2.1 11βHSD2

The 11βHSD2 experimental data, used for parameterisation, were collected from transfected cells with only 11βHSD2 present. Varying initial concentrations of either 11OHA4, 11β-Hydroxy-5α-Dione (11OH5αDione), or 11β-Hydroxyandrosterone (11OHAST) were converted to the products 11KA4, 11K5αDione, or 11-Keto-androsterone (11KAST) respectively. The initial concentrations were 0.25, 0.5, 0.75, 1.0, 2.5, and 5.0 μM, and all experiments were performed in triplicate.

6.2.1.1 Parameter estimation

Irreversible Michaelis-Menten kinetics were assumed and the rate equations for the reactions catalysed by 11βHSD2 are shown as equations 6.1 to 6.3. The conversion of 11OHA4 to 11KA4 is reaction v1, the conversion of 11OH5αDione to 11K5αDione is reaction v2, and the conversion of 11OHAST to 11KAST is reaction v3. The ordinary differential equations 6.4a to 6.4f show the change in metabolite concentrations over time. This ODE based model is fitted to the experimental data and the respective V_{max} and K_m values are determined by minimising the objective function 6.5. The function `minimizeObjectiveFunction` of the package `IdentifiabilityAnalysis` was used for model fitting. All experiments were performed in triplicate and the objective function is weighted to the variance in the experimental data. Table 6.4 shows the parameter estimations with confidence intervals for the enzyme 11βHSD2. Figure 6.3 shows the experimental data and model fits.

Rate equations

$$v_1 = \frac{Vmax_{11OHA4} \cdot [11OHA4]}{Km_{11OHA4} + [11OHA4]} \tag{6.1}$$

$$v_2 = \frac{Vmax_{11OH5aDione} \cdot [11OH5aDione]}{Km_{11OH5aDione} + [11OH5aDione]} \tag{6.2}$$

$$v_3 = \frac{Vmax_{11OHAST} \cdot [11OHAST]}{Km_{11OHAST} + [11OHAST]} \tag{6.3}$$

Ordinary differential equations

$$\frac{d}{dt}[11OHA4] = -v_1 \tag{6.4a}$$

$$\frac{d}{dt}[11OH5\alpha Dione] = -v_2 \tag{6.4b}$$

$$\frac{d}{dt}[11OHAST] = -v_3 \tag{6.4c}$$

$$\frac{d}{dt}[11KA4] = v_1 \tag{6.4d}$$

$$\frac{d}{dt}[11K5\alpha Dione] = v_2 \tag{6.4e}$$

$$\frac{d}{dt}[11KAST] = v_3 \tag{6.4f}$$

Objective function

$$SSR = \sum_{i=1}^{n} \frac{(data_i - model_i)^2}{\sigma^2} \tag{6.5}$$

where n is the number of experimental datasets, $data_i$ is the experimental data, $model_i$ is the model fit to $data_i$ and σ_i^2 is the variance in the experimental data.

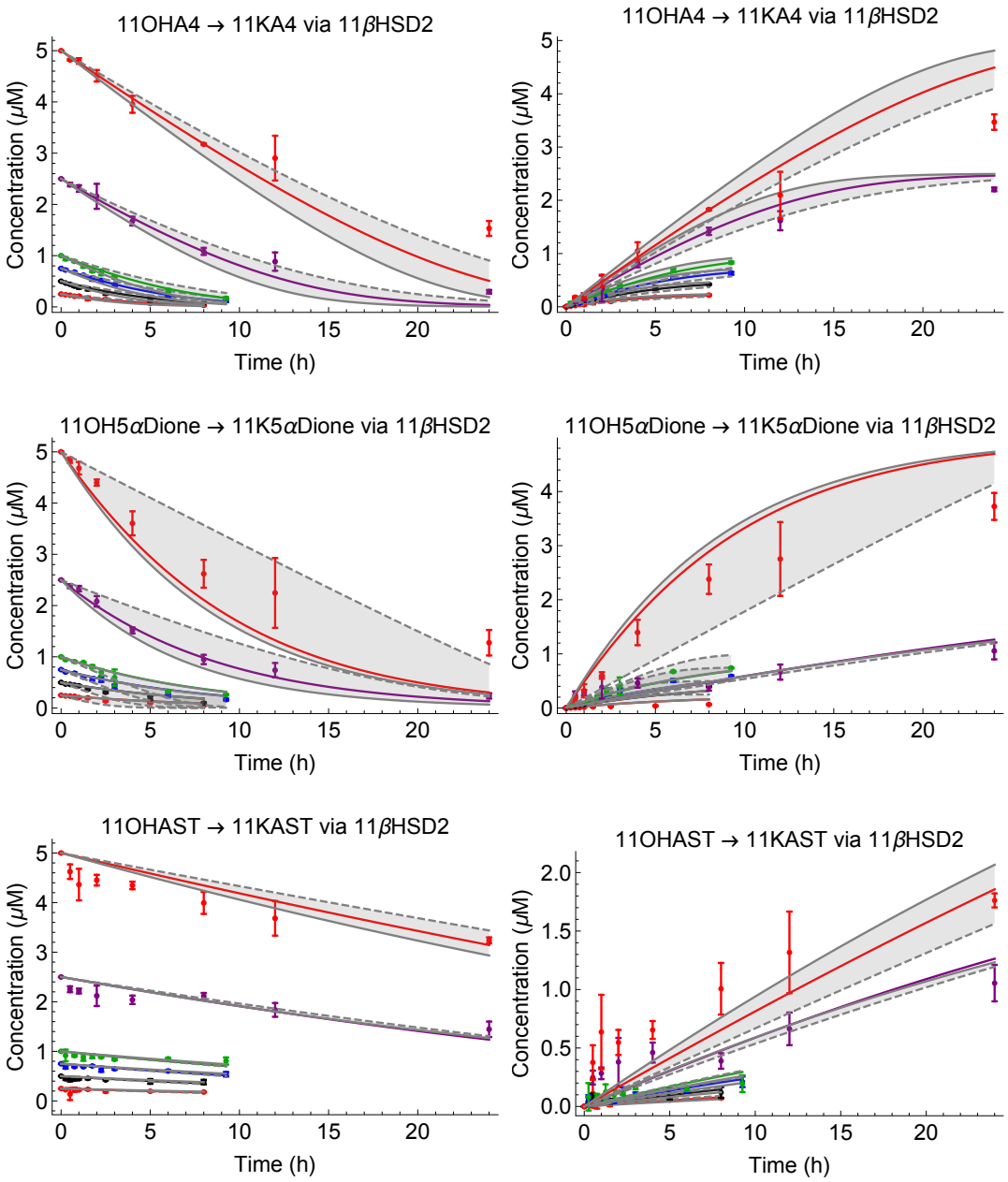


Figure 6.3: **Model fit (solid lines) with 95% confidence intervals (grey areas) to experimental 11βHSD2 data (points).** The experiments were performed in triplicate, with the variance indicated by the error bars.

Table 6.4: **Fitted parameter values for 11βHSD2 kinetics** with their respective 95% confidence intervals shown in brackets.

Substrate	V_{max} (μM/h)	K_m (μM)
11OHA4	0.286 (0.264 - 0.311)	1.031 (0.810 - 1.31)
11OH5αDione	2.342 (0.307 - 4.684)	18.32 (1.98 - 36.64)
11OHAST	0.125 (0.086 - 0.213)	2.476 (1.34 - 5.71)

6.2.1.2 Identifiability analysis

The identifiability of all the newly determined parameter values were tested. The log-likelihood ratio test was used to determine the 95% confidence intervals as well as whether the parameter values are identifiable. The functions `profileLikelihood` and `likelihoodRatioPlot` of the package `IdentifiabilityAnalysis` were used to test the identifiability of the parameters and create the profile likelihood plots. All of the estimated values were identifiable except for the V_{max} and K_m values for the conversion of 11OH5αDione to 11K5αDione. These two values were practically non-identifiable, with a 95% confidence interval that is lower bound, but not upper bound. With the experimental data used for model parameterisation, the substrate concentrations were not high enough to saturate the enzyme, therefore there is no upper limit to the confidence interval. In the analysis shown above, 2 times the parameter value is used for the upper limit of the confidence interval if the parameter value was not identifiable. All the profile likelihood plots of the identifiability analysis of the parameters are shown in Figure 6.4.

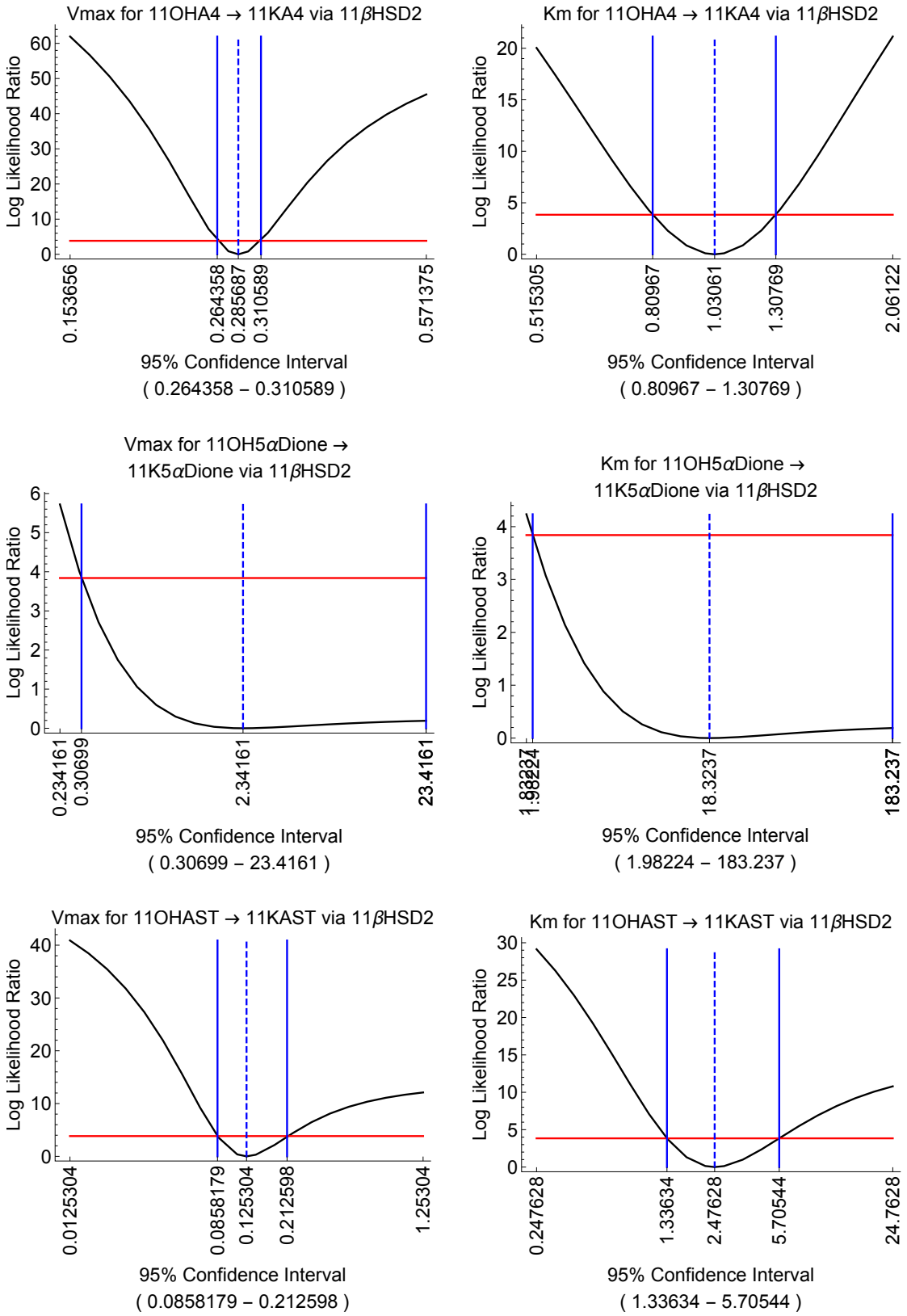


Figure 6.4: **Profile likelihood plots of the 11βHSD2 parameters.** The solid blue lines are the 95% confidence intervals that constrain the fitted parameter value, indicated by the dashed blue line. The red lines indicate the 95% confidence threshold value.

6.2.1.3 Validation

The parameter value estimations were validated by using these values to predict the outcome of run-out experiments of another transfection. To account for the differences in transfection efficiency between the two sets of data, the initial rate of conversion of 1 µM 11OH5αDione was compared between the two experiments. The ratio of the difference in the initial rates were used as a variable multiplier while predicting the outcome of the run-out experiment for the conversion of 11OH5αDione to 11K5αDione. With this variable multiplier for 1 µM 11OH5αDione conversion, the run-out experiment of 1 µM 11OHA4 to 11KA4 was also predicted as validation. The model predictions and the run out data are shown in Figure 6.5.

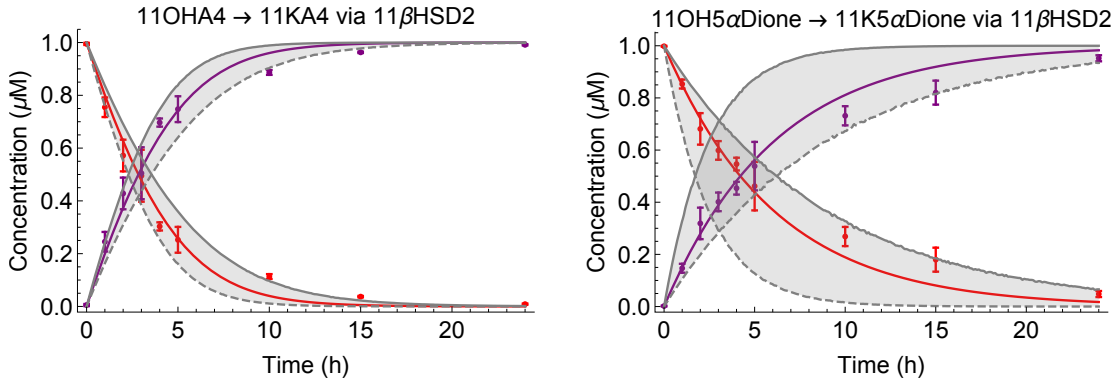


Figure 6.5: **Model validation: prediction (solid lines) of 11βHSD2 run-out data (points). The 95% confidence intervals of the model prediction is shown by the grey ares.**

6.2.2 Newly determined parameter values: 3αHSD

To the best of our knowledge, no kinetic data has been collected describing the dynamics of 3αHSD. Mass action kinetics for this enzyme are assumed and for each 3αHSD reaction only one parameter value is fitted for, which describes the ratio of V_{max}/K_m. These parameter values are fitted for during model construction with the use of C42B cancer cell data as this is the only data available that includes 3αHSD kinetics. This is further addressed in the following section.

6.3 Model construction and validation

The V_{max} and K_m pairs of all the enzymes (except 3αHSD) have been determined. These parameter value pairs have been determined in transfected cells. These values are then combined to construct a complete model for steroid biosynthesis in prostate cancer cells. The transfection efficiency of each enzyme is different than the expression level of

the enzymes in the cancer cells. To adjust for the differences in enzyme expression, a variable multiplier is fitted for for each of the enzymes. Data collected from prostate cancer cell line C42B were used to construct the model. This cell line expresses the enzymes AKR1C3, 17βHSD2, SRD5A, 11βHSD2, and 3αHSD. Three variable multipliers were fitted for: one each for SRD5A and 11βHSD2, and the third multiplier adjusted for the expression of both AKR1C3 and 17βHSD2. The ratio of AKR1C3 to 17βHSD2 expression, as published in [7], was included in the fitting and therefore it was possible to fit for one less multiplier. These variable multiplier values are incorporated into the model to ensure that the model describes the correct relative activities of the enzymes.

6.3.1 3αHSD kinetic parameters

Mass action kinetics for this enzyme are assumed and only one parameter value is fitted for, for each 3αHSD reaction. These mass action parameter values were fitted for along with the enzyme expression variable multipliers. The same parameter estimation technique of nonlinear regression and minimising of the objective function 6.5 was used for 3αHSD. The mass action kinetic values $k_{5aDione}$, k_{5aDHT}, k_{AST}, and $k_{3aAdiol}$, $k_{11OH5aDione}$, $k_{11K5aDione}$, $k_{11K5aDHT}$, $k_{11OHAST}$, and k_{11KAST} were fitted for.

6.3.2 C42B data for model construction and validation

Time course datasets from three experiments in C42B cells were collected. To construct the model, the data of one of the experiments were used to fit for all variable multipliers and 3αHSD parameters. As validation, the model is then used to predict the outcome of the other two datasets. Two datasets were used for validation and only one was used for model construction as two of the datasets were biological repeats. Of the two datasets that are biological repeats, one was used for construction and the other was used for validation. The third dataset was therefore kept as a validation dataset, otherwise the only dataset used for validation would have been one that was nearly identical to a dataset used for model construction.

6.3.3 Model construction

The dataset used for model construction consists of four time courses from one transfection. The data is indicated as Transfection 1 in Table 6.5. Figure 6.6 shows the model fit to the model construction data. During this fit, the 38 previously determined parameter values of AKR1C3, 17βHSD2, SRD5A, and 11βHSD2 were kept as is. The three variable multipliers (α, β and γ), for the relative expression of these enzymes,

Table 6.5: **C42B data** for model construction and validation

Dataset	Substrate concentration	Products measured
Transfection 1	0.1 uM 11OHA4	All metabolites in Fig. 6.2
	1.0 uM 11OHA4	All metabolites in Fig. 6.2
	0.1 uM A4	All metabolites in Fig. 6.1
	1.0 uM A4	All metabolites in Fig. 6.1
Transfection 2	1.0 uM 11OHA4	11KA4, 11KT, 11K5αDHT, 11KAST
	1.0 uM A4	T, 5αDione, 5αDHT
Transfection 3	0.1 11OHA4	All metabolites in Fig. 6.2
	1.0 11OHA4	All metabolites in Fig. 6.2
	0.1 A4	All metabolites in Fig. 6.1
	1.0 A4	All metabolites in Fig. 6.1

were fitted for, along with the nine mass action parameter values for the nine reactions catalysed by 3αHSD (see rate equations 6.6a to 6.6l and equations 6.7a to 6.7p). The ODEs (equation 6.9a to 6.9l) were fitted to a single set of experimental data and the 12 unknowns were determined by minimising an objective function with weighted nonlinear least squares regression. A total of 12 unknown values were fitted for with a single dataset, consisting of four time courses. Prior to running a full identifiability analysis on the 3αHSD parameter values, these values were varied with a factor 10 smaller and larger than the estimated value. These changes did not affect the variables in the top two tiers of the pathways in Figures 6.1 and 6.2, indicating that these parameter values are not identifiable. Despite the limited amount of data available for model construction and the non-identifiability of the 3αHSD parameter values, a fairly accurate fit was obtained, as can be seen in Figure 6.6. The fitted values of these three multipliers were incorporated into the previously determined parameter values for AKR1C3, 17βHSD2, SRD5A, and 11βHSD2. These updated parameter values, which now include the relative enzyme activities, and the newly determined 3αHSD values are shown in Tables 6.6 and 6.7. The parameter values determined for 3αHSD were used as is, despite their non-identifiability, as this did not affect the success of the model in terms of describing the dynamics of the metabolite synthesis of the top two tiers of the pathways in Figures 6.1 and 6.2

Rate equations

$$v_1 = \frac{\alpha \cdot Vmax^{A4}_{AKR1C3} \cdot [A4]/Km^{A4}_{AKR1C3}}{denom_{AKR1C3}} \tag{6.6a}$$

$$v_2 = \frac{\alpha \cdot Vmax^{5aDione}_{AKR1C3} \cdot [5\alpha Dione]/Km^{5aDione}_{AKR1C3}}{denom_{AKR1C3}} \tag{6.6b}$$

$$v_3 = \frac{\alpha \cdot Vmax^{AST}_{AKR1C3} \cdot [AST]/Km^{AST}_{AKR1C3}}{denom_{AKR1C3}} \tag{6.6c}$$

$$v_4 = \frac{\alpha \cdot Vmax^{T}_{17BHSD2} \cdot [T]/Km^{T}_{17BHSD2}}{denom_{17BHSD2}} \tag{6.6d}$$

$$v_5 = \frac{\alpha \cdot Vmax^{5aDHT}_{17BHSD2} \cdot [5\alpha DHT]/Km^{5aDHT}_{17BHSD2}}{denom_{17BHSD2}} \tag{6.6e}$$

$$v_6 = \frac{\alpha \cdot Vmax^{3aAdiol}_{17BHSD2} \cdot [3\alpha Adiol]/Km^{3aAdiol}_{17BHSD2}}{denom_{17BHSD2}} \tag{6.6f}$$

$$v_7 = \frac{\beta \cdot Vmax^{A4}_{SRD5A} \cdot [A4]/Km^{A4}_{SRD5A}}{denom_{SRD5A}} \tag{6.6g}$$

$$v_8 = \frac{\beta \cdot Vmax^{T}_{SRD5A} \cdot [T]/Km^{T}_{SRD5A}}{denom_{SRD5A}} \tag{6.6h}$$

$$v_9 = k^{5aDione}_{3aHSD} \cdot [5\alpha Dione] \tag{6.6i}$$

$$v_{10} = k^{5aDHT}_{3aHSD} \cdot [5\alpha DHT] \tag{6.6j}$$

$$v_{11} = k^{AST}_{3aHSD} \cdot [AST] \tag{6.6k}$$

$$v_{12} = k^{3aAdiol}_{3aHSD} \cdot [3\alpha Adiol] \tag{6.6l}$$

Rate equations

$$v_{13} = \frac{\gamma \cdot Vmax_{11BHSD2}^{11OHA4} \cdot [11OHA4]/Km_{11BHSD2}^{11OHA4}}{denom_{11BHSD2}} \tag{6.7a}$$

$$v_{14} = \frac{\gamma \cdot Vmax_{11BHSD2}^{11OH5aDione} \cdot [11OH5\alpha Dione]/Km_{11BHSD2}^{11OH5aDione}}{denom_{11BHSD2}} \tag{6.7b}$$

$$v_{15} = \frac{\gamma \cdot Vmax_{11BHSD2}^{11OHAST} \cdot [11OHAST]/Km_{11BHSD2}^{11OHAST}}{denom_{11BHSD2}} \tag{6.7c}$$

$$v_{16} = \frac{\alpha \cdot Vmax_{AKR1C3}^{11KA4} \cdot [11KA4]/Km_{AKR1C3}^{11KA4}}{denom_{AKR1C3}} \tag{6.7d}$$

$$v_{17} = \frac{\alpha \cdot Vmax_{AKR1C3}^{11K5aDione} \cdot [11K5\alpha Dione]/Km_{AKR1C3}^{11K5aDione}}{denom_{AKR1C3}} \tag{6.7e}$$

$$v_{18} = \frac{\alpha \cdot Vmax_{AKR1C3}^{11KAST} \cdot [11KAST]/Km_{AKR1C3}^{11KAST}}{denom_{AKR1C3}} \tag{6.7f}$$

$$v_{19} = \frac{\alpha \cdot Vmax_{17BHSD2}^{11KT} \cdot [11KT]/Km_{17BHSD2}^{11KT}}{denom_{11KT}} \tag{6.7g}$$

$$v_{20} = \frac{\alpha \cdot Vmax_{17BHSD2}^{11K5aDHT} \cdot [11K5\alpha DHT]/Km_{17BHSD2}^{11K5aDHT}}{denom_{17BHSD2}} \tag{6.7h}$$

$$v_{21} = \frac{\beta \cdot Vmax_{SRD5A}^{11OHA4} \cdot [11OHA4]/Km_{SRD5A}^{11OHA4}}{denom_{SRD5A}} \tag{6.7i}$$

$$v_{22} = \frac{\beta \cdot Vmax_{SRD5A}^{11KA4} \cdot [11KA4]/Km_{SRD5A}^{11KA4}}{denom_{SRD5A}} \tag{6.7j}$$

$$v_{23} = \frac{\beta \cdot Vmax_{SRD5A}^{11KT} \cdot [11KT]/Km_{SRD5A}^{11KT}}{denom_{SRD5A}} \tag{6.7k}$$

$$v_{24} = k_{3aHSD}^{11OH5aDione} \cdot [11OH5\alpha Dione] \tag{6.7l}$$

$$v_{25} = k_{3aHSD}^{11K5aDione} \cdot [11K5\alpha Dione] \tag{6.7m}$$

$$v_{26} = k_{3aHSD}^{11K5aDHT} \cdot [11K5\alpha DHT] \tag{6.7n}$$

$$v_{27} = k_{3aHSD}^{11OHAST} \cdot [11OHAST] \tag{6.7o}$$

$$v_{28} = k_{3aHSD}^{11KAST} \cdot [11KAST] \tag{6.7p}$$

$$denom_{AKR1C3} = 1 + \frac{A4}{Km^{A4}_{AKR1C3}} + \frac{5\alpha Dione}{Km^{5aDione}_{AKR1C3}} + \frac{AST}{Km^{AST}_{AKR1C3}} + \frac{11KA4}{Km^{11KA4}_{AKR1C3}} + \frac{11K5\alpha Dione}{Km^{11K5aDione}_{AKR1C3}} + \frac{11KAST}{Km^{11KAST}_{AKR1C3}}$$

$$denom_{17BHSD2} = 1 + \frac{T}{Km^{T}_{17BHSD2}} + \frac{5\alpha DHT}{Km^{5aDHT}_{17BHSD2}} + \frac{3\alpha Adiol}{Km^{3aAdiol}_{17BHSD2}} + \frac{11KT}{Km^{11KT}_{17BHSD2}} + \frac{11K5\alpha DHT}{Km^{11K5aDHT}_{17BHSD2}}$$

$$denom_{SRD5A} = 1 + \frac{A4}{Km^{A4}_{SRD5A}} + \frac{T}{Km^{T}_{SRD5A}} + \frac{11KA4}{Km^{11KA4}_{SRD5A}} + \frac{11KT}{Km^{11KT}_{SRD5A}} + \frac{11OHA4}{Km^{11OHA4}_{SRD5A}}$$

$$denom_{11BHSD2} = 1 + \frac{11OHA4}{Km^{11OHA4}_{11BHSD2}} + \frac{11OH5\alpha Dione}{Km^{11OH5aDione}_{11BHSD2}} + \frac{11OHAST}{Km^{11OHAST}_{11BHSD2}}$$

Ordinary differential equations

$$\frac{d}{dt}[T] = -v_1 - v_4 - v_8 \tag{6.9a}$$

$$\frac{d}{dt}[5\alpha DHT] = v_2 - v_5 + v_8 - v_{10} + v_{12} \tag{6.9b}$$

$$\frac{d}{dt}[5\alpha Dione] = -v_2 + v_5 + v_7 - v_9 + v_{11} \tag{6.9c}$$

$$\frac{d}{dt}[AST] = -v_3 + v_6 + v_9 - v_{11} \tag{6.9d}$$

$$\frac{d}{dt}[3\alpha Adiol] = v_3 - v_6 + v_{10} - v_{12} \tag{6.9e}$$

$$\frac{d}{dt}[11OH5\alpha Dione] = -v_{14} + v_{21} - v_{24} + v_{27} \tag{6.9f}$$

$$\frac{d}{dt}[11OHAST] = -v_{15} + v_{24} - v_{27} \tag{6.9g}$$

$$\frac{d}{dt}[11KA4] = v_{13} - v_{16} + v_{19} - v_{22} \tag{6.9h}$$

$$\frac{d}{dt}[11K5\alpha Dione] = v_{14} - v_{17} + v_{20} + v_{22} - v_{25} + v_{28} \tag{6.9i}$$

$$\frac{d}{dt}[11KAST] = v_{15} - v_{18} + v_{25} - v_{28} \tag{6.9j}$$

$$\frac{d}{dt}[11KT] = v_{16} - v_{19} - v_{23} \tag{6.9k}$$

$$\frac{d}{dt}[11K5\alpha DHT] = v_{17} - v_{20} + v_{23} - v_{26} \tag{6.9l}$$

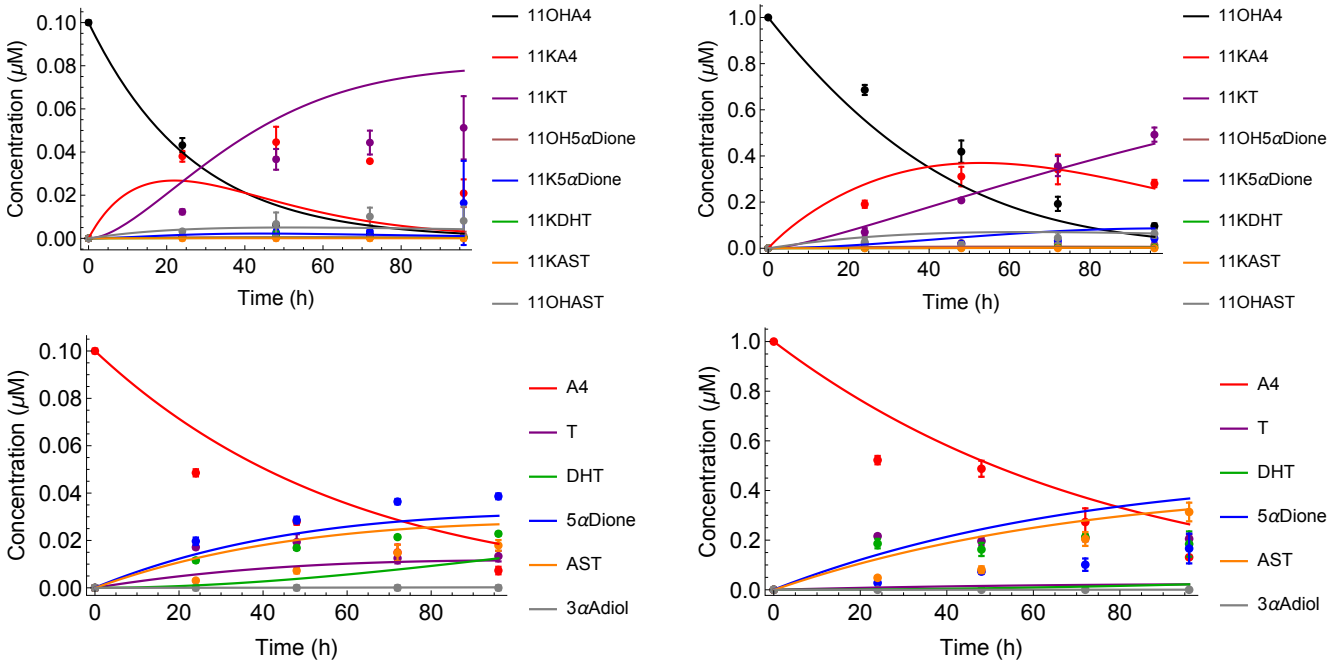


Figure 6.6: **Model construction: fitting of one C42B dataset to determine relative enzyme expression and all 3αHSD mass action parameter values.**

Table 6.6: **Classic androgen parameter values in C42B CRPC**

Parameter	μM/h	Parameter	μM
$V_{max}^{A4}{}_{AKR1C3}$	0.0005	$K_m^{A4}{}_{AKR1C3}$	0.06
$V_{max}^{5aDione}{}_{AKR1C3}$	0.0004	$K_m^{5aDione}{}_{AKR1C3}$	0.12
$V_{max}^{AST}{}_{AKR1C3}$	0.0014	$K_m^{AST}{}_{AKR1C3}$	0.27
$V_{max}^{T}{}_{17BHSD2}$	0.0019	$K_m^{T}{}_{17BHSD2}$	5.08
$V_{max}^{5aDHT}{}_{17BHSD2}$	0.0012	$K_m^{5aDHT}{}_{17BHSD2}$	3.77
$V_{max}^{3aAdiol}{}_{17BHSD2}$	0.0021	$K_m^{3aAdiol}{}_{17BHSD2}$	5.86
$V_{max}^{A4}{}_{SRD5A}$	0.1681	$K_m^{A4}{}_{SRD5A}$	12.1
$V_{max}^{T}{}_{SRD5A}$	2.384	$K_m^{T}{}_{SRD5A}$	373.0
$k_{3aHSD}^{5aDione}$	4.096	k_{3aHSD}^{5aDHT}	0.1660
k_{3aHSD}^{AST}	4.639	$k_{3aHSD}^{3aAdiol}$	10.29

Table 6.7: **11-Oxygenated androgen parameter values in C42B CRPC cells**

Parameter	μM/h	Parameter	μM
$V_{max}^{11OHA4\ 11BHSD2}$	0.0386	$K_m^{11OHA4\ 11BHSD2}$	1.03
$V_{max}^{11OH5aDione\ 11BHSD2}$	0.3111	$K_m^{11OH5aDione\ 11BHSD2}$	18.3
$V_{max}^{11OHAST\ 11BHSD2}$	0.0166	$K_m^{11OHAST\ 11BHSD2}$	2.48
$V_{max}^{11KA4\ AKR1C3}$	0.0074	$K_m^{11KA4\ AKR1C3}$	0.11
$V_{max}^{11K5aDione\ AKR1C3}$	0.0151	$K_m^{11K5aDione\ AKR1C3}$	0.23
$V_{max}^{11KAST\ AKR1C3}$	0.0162	$K_m^{11KAST\ AKR1C3}$	0.34
$V_{max}^{11KT\ 17BHSD2}$	0.0022	$K_m^{11KT\ 17BHSD2}$	9.11
$V_{max}^{11K5aDHT\ 17BHSD2}$	0.0033	$K_m^{11K5aDHT\ 17BHSD2}$	7.72
$V_{max}^{11OHA4\ SRD5A}$	0.0963	$K_m^{11OHA4\ SRD5A}$	30.4
$V_{max}^{11KA4\ SRD5A}$	0.3973	$K_m^{11KA4\ SRD5A}$	90.7
$V_{max}^{11KT\ SRD5A}$	0.5348	$K_m^{11KT\ SRD5A}$	918
$k_{3aHSD}^{11OH5aDione}$	4.820	$k_{3aHSD}^{11K5aDione}$	0.2918
$k_{3aHSD}^{11K5aDHT}$	0.7557	$k_{3aHSD}^{11OHAST}$	0.4855
k_{3aHSD}^{11KAST}	14.09		

6.3.4 Model validation

The remaining two C42B datasets were used for model validation. As mentioned above, the experimental conditions of the one experiment were the same as the experimental conditions of the model construction dataset i.e. these two datasets are biological repeats. This dataset consists of four time courses and is indicated as Transfection 3 in Table 6.5. The prediction of these experimental results are shown in Figure 6.7. The other experiment consists of two time courses and is indicated as Transfection 2 in Table 6.5. The model prediction of this experiment is shown in Figure 6.8.

Fig. 6.7A shows that the model overestimates the conversion of 11KA4 to 11KT with an initial concentration of 0.1 μM 11OHA4. Fig. 6.7B shows that at the higher initial concentration of 1.0 μM 11OHA4 the model gives an accurate prediction of the

experimental data. At the lower initial concentration of 0.1 µM A4 (Fig. 6.7C), the model gives an qualitative prediction of the experimental results, however at the higher initial concentration of 1.0 µM A4 (Fig. 6.7D) the model returns an underestimation of the synthesis of 5αDHT. Similar results are seen in Fig. 6.8B for the same initial concentration of A4, the model underestimates the synthesis of 5αDHT. In this case the consumption of A4 is also underestimated. In Fig. 6.8A, with an initial concentration of 1.0 11OHA4, the consumption of 11OHA4, and the synthesis of intermediates and products are underestimated.

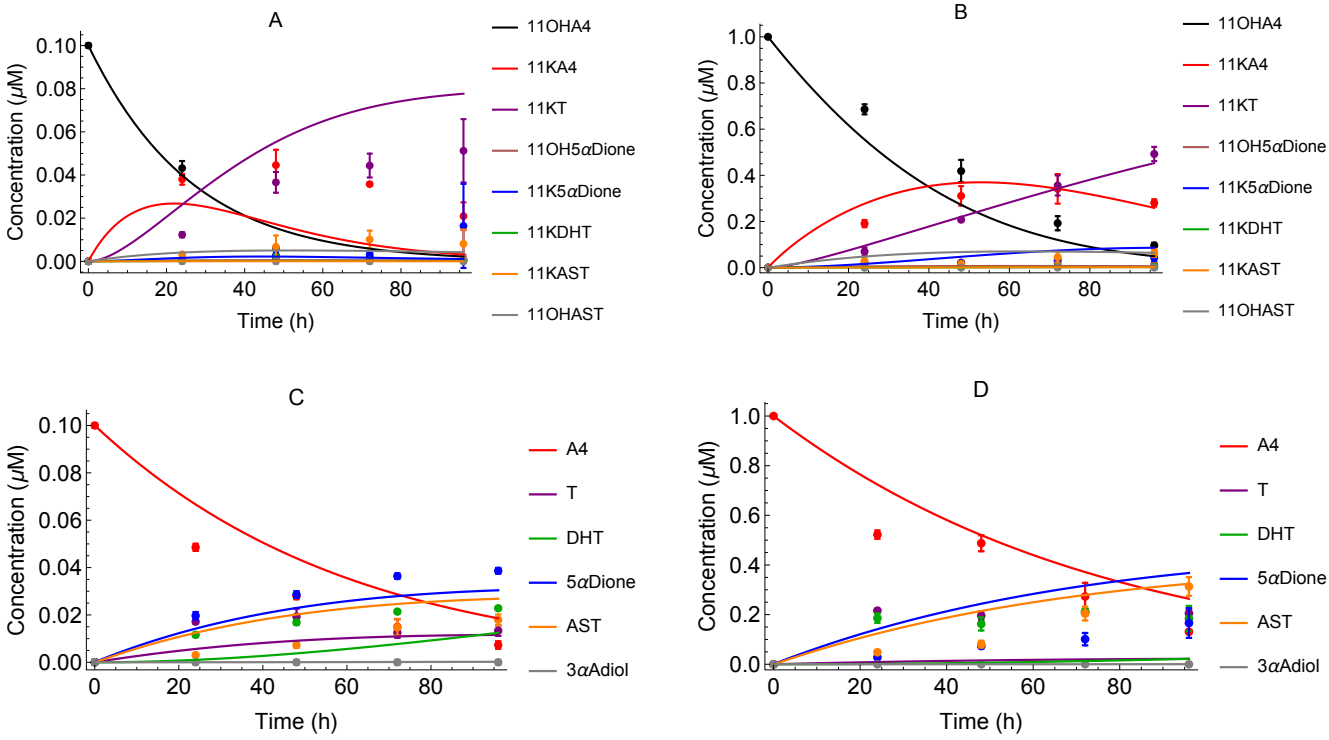


Figure 6.7: **Model validation: predicting the results of C42B datasets (Transfection 3).**

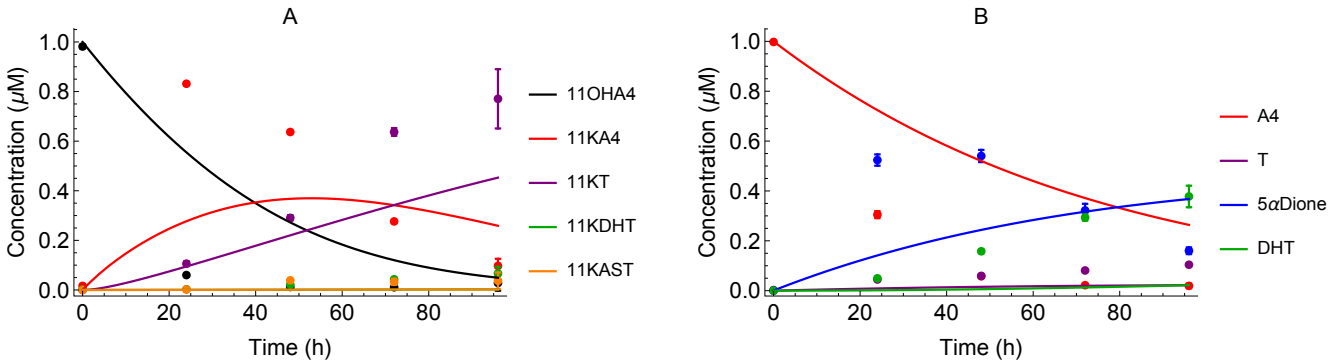


Figure 6.8: **Model validation: predicting the results of C42B datasets (Transfection 2).**

6.4 Identifying possible treatment targets

The following section contains steady state and metabolic control analysis (MCA) of the CRPC model in C42B cells. Since we need a steady state for the MCA, we first adapt the model so that it can reach steady state.

6.4.1 Steady state analysis

Figure 6.9 shows the steady state concentrations of all the steroid intermediates and products with fixed concentrations of 3.9 nM A4 and 3.7 nM 11OHA4. These concentrations represent the physiological serum concentrations of A4 and 11OHA4 in men [94]. All intermediates and products are removed from the system at a constant rate of 0.1 h^{-1} times the metabolite concentration. This technique of steady state analysis of a steroid hormone model is similar to the method used by Nguyen *et al* in [81]. According to the model simulations, when an equal supply of A4 and 11OHA4 is present in the cell, 11KA4 and 11KT have the highest steady state concentrations.

These results are robust. Varying the supply of A4 and 11OHA4 returns nearly the exact same results. Nguyen *et al* [81] suggests a removal value of metabolite intermediates and products between 0.01 and 1.0 $time^{-1}$, they however use the value of 0.1 $time^{-1}$ in their analysis. Changing the removal of metabolite intermediates and products to 0.01 h^{-1} shows that 11KT has the highest steady state value. Increasing the removal value to 1.0 h^{-1} however has similar results to Fig. 6.9.

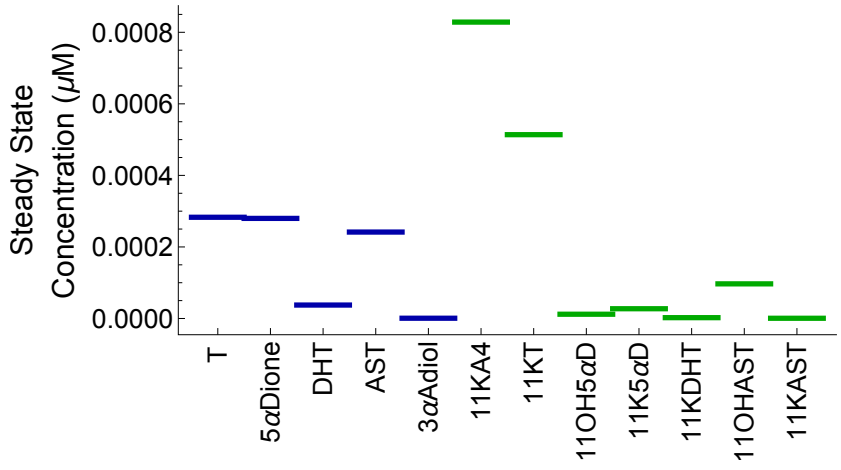


Figure 6.9: **Steady state concentrations (μM) of classic androgens (blue) and 11-oxygenated androgens (green)** with a supply of 10 nM 11OHA4 and A4 per hour and removal of all intermediates and products at a rate of 0.1 h^{-1} times the concentration.

6.4.2 Flux control and concentration control

The following section addresses the flux control and concentration control of the enzymes, this is shown in the heat maps of Figures 6.12 and 6.13. The fluxes are numbered from J1 to J28. Figures 6.10 and 6.11 show which flux corresponds to which branch in either the classic androgen or 11-oxygenated androgen synthesis pathway.

6.4.2.1 3αHSD

The enzyme 3αHSD has strong positive control over its own reactions, except the reaction J26 over which it has a less strong control. The only reaction, other than its own, that it has any significant effect on is the reaction J20 (the conversion of 11K5αDHT to 11K5αDione by 17βHSD2), see Fig. 6.12.

3αHSD has very little control over any metabolite concentrations, except for a strong negative control over the concentration of 11KAST. This enzyme has little control over the processes leading to the potent androgens T, 5αDHT, 11KT, and 11K5αDHT, see Fig. 6.13.

6.4.2.2 11βHSD2

This enzyme has strong positive control over nearly every reaction in the 11-oxygenated pathway. Not only does it have control over it's own reactions (J13, J14, and J15), but it has positive control over all AKR1C3 and 17βHSD2's reactions (J16, J17, and J18 and J19 and J20 respectively) in the 11-oxygenated pathway. It also has control over two of SRD5A's reactions (J22 and J23), and three of 3αHSD's reactions (J25, J26, and J28), see Fig. 6.12.

11βHSD2 has little control over any of the metabolites in the top tier of the 11-oxygenated pathway (11KA4 or 11KT), but it has a strong positive control over the concentrations of all the other metabolites in this pathway, see Fig. 6.13. 11βHSD2 has virtually no control over any of the reactions or metabolite concentrations of the classic androgen pathway.

6.4.2.3 SRD5A

This enzyme has most control over the system in both the classic and 11-oxygenated androgen pathways. It has strong positive control over all the reactions except for J1 and J4 in the classic pathway, and J13, J16, and J19 in the 11-oxygenated pathway. These reactions are all in the top tiers of the pathways. SRD5A has strong positive

control over all the reactions in the middle and bottom tiers of the pathways, see Fig. 6.12.

SRD5A has insignificant control over the concentration of T in the classic pathway, but has strong positive control over all metabolites in the middle and bottom tiers of the classic pathway. In contrast, although SRD5A has little control over the reactions of the top tier of the 11-oxygenated pathway, it has strong positive control over the concentrations of the metabolites in the top tier (11KA4 and 11KT). SRD5A has strong positive control over all metabolite concentrations of the middle and bottom tiers of the 11-oxygenated pathway, except for 11OH5αDione and 11OHAST, see Fig. 6.13.

6.4.2.4 17βHSD2

This enzyme has strong positive control over it's own reactions in both pathways (J4, J5, J6, J19, and J20), but has little control over any other reaction, see Fig. 6.12. 17βHSD2 also has insignificant control over any metabolite concentration, see Fig. 6.13.

6.4.2.5 AKR1C3

In the classic androgen pathway, AKR1C3 has strong positive control over it's own reactions and all the reactions catalysed by 17βHSD2 (J1, J2, and J3, and J4, J5, and J6 respectively), as can be seen in Figure 6.12. It also has strong positive control over all the reactions downstream of itself, i.e reaction J8 catalysed by SRD5A and reactions J10 and J12 catalysed by 3αHSD. In contrast, AKR1C3 does not have have very strong control over any of the reactions in the 11-oxygenated pathway. It has moderate positive control over all reactions catalysed by itself and 17βHSD2 (J16, J17, and J18 and J19 and J20 respectively). AKR1C3 has moderate negative control over the reactions running from 11KA4 to and between 11K5αDione and 11KAST (J22, J25, and J28). It has moderate positive control over the reactions running from 11KT to 11K5αDHT and the further conversion of 11KDHT (J23 and J26), see Fig. 6.12.

In the classic androgen pathway, AKR1C3 has strong positive control over the concentrations of T, 5αDione, and 3αAdiol. In the 11-oxygenated pathway, AKR1C3 has moderate negative control over the concentrations of the metabolites of the middle tier (11OH5αDione, 11K5αDione, and 11KDHT) and moderate positive control over the concentrations of the metabolites of the bottom tier (11OHAST and 11KAST), see Fig. 6.13.

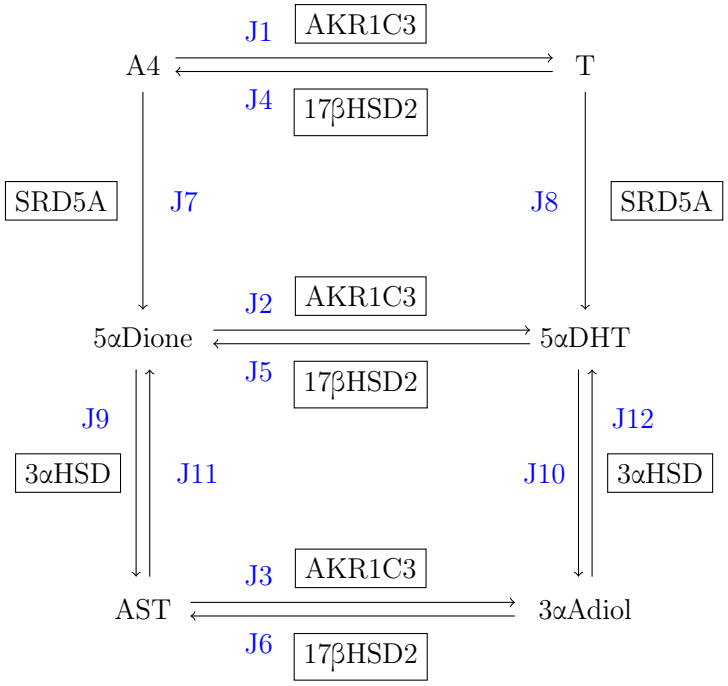


Figure 6.10: **Schematic representation of the classic androgen and 5αDione biosynthesis pathway.** The fluxes through the branches are numbered from J1 to J12.

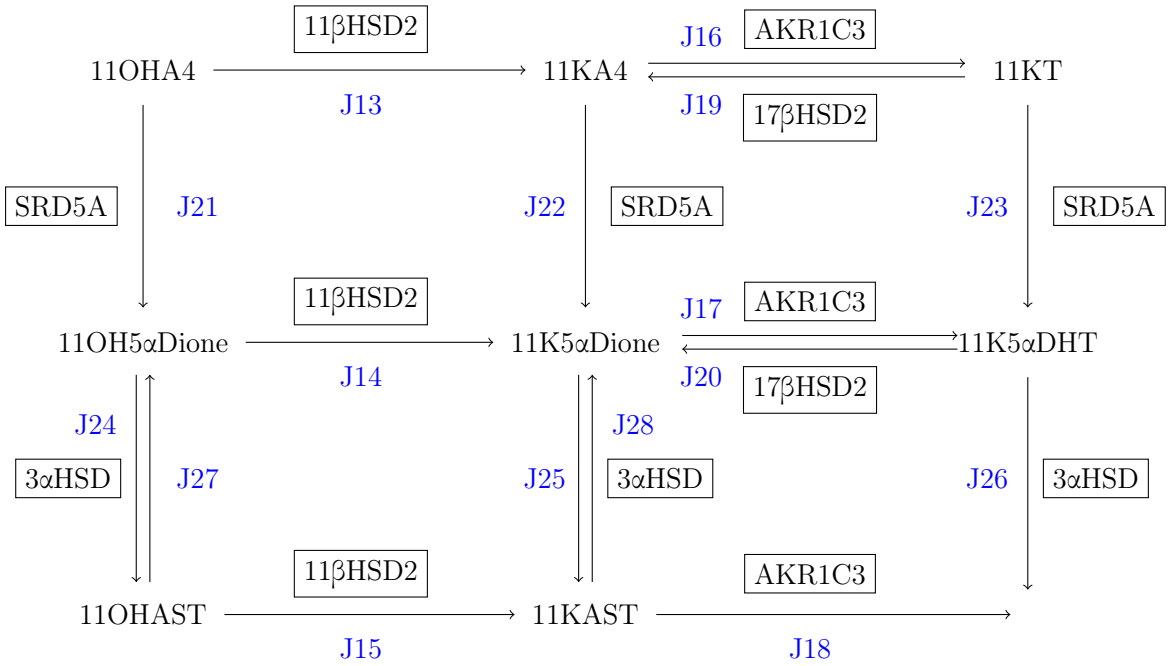


Figure 6.11: **Schematic representation of the 11-oxygenated androgen biosynthesis pathway.** The fluxes through the pathway are numbered from J13 to J28.

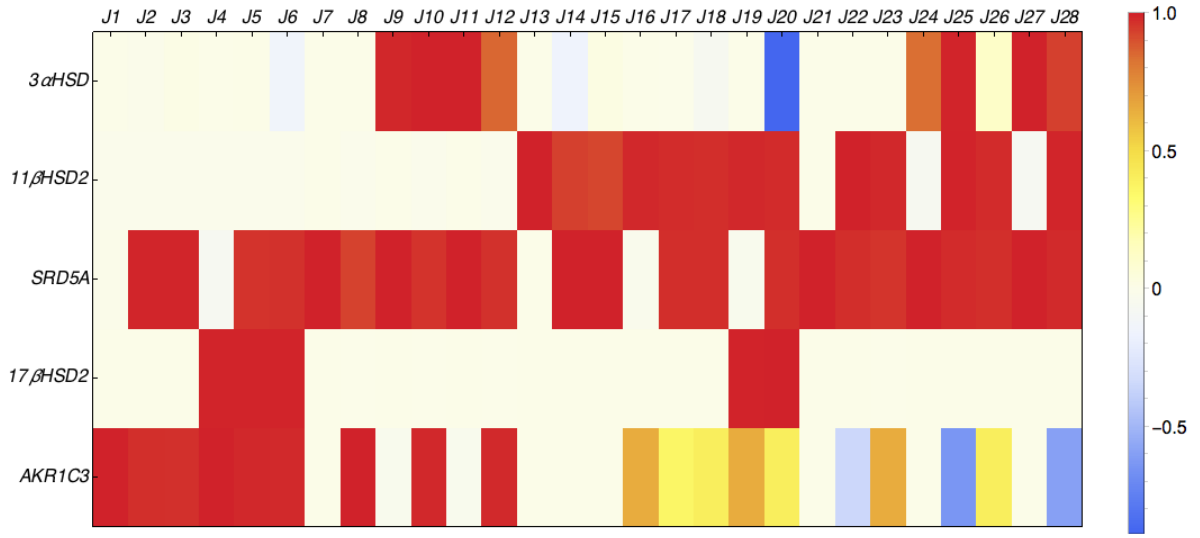


Figure 6.12: **Heat map of the flux control coefficients of the enzymes AKR1C3, 17βHSD2, SRD5A, 11βHSD2, and 3αHSD on the flux through the branches J1 to J28** (see Fig. 6.10 and 6.11).

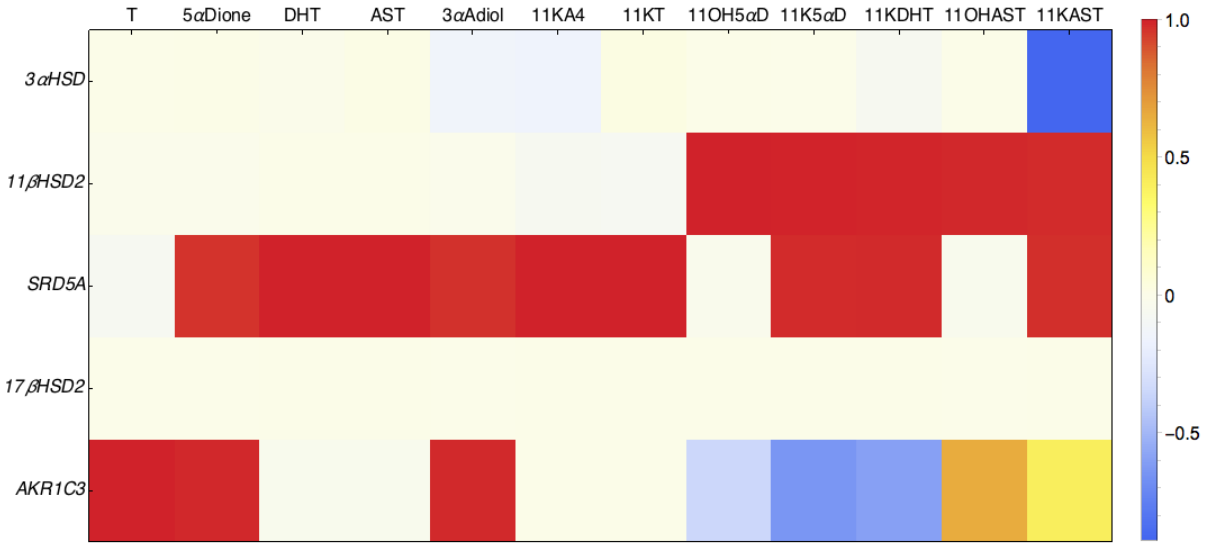


Figure 6.13: **Heat map of the concentration control coefficients of the enzymes AKR1C3, 17βHSD2, SRD5A, 11βHSD2, and 3αHSD on the concentrations of the metabolites** T, 5αDione, 5αDHT, AST, 3αAdiol, 11KA4, 11KT, 11OH5αDione, 11K5αDione, 11K5αDHT, 11OHAST, and 11KAST.

The MCA shows that the three enzymes with the most control over the flux through the system as well as the most control over the majority of the metabolite concentrations are AKR1C3, SRD5A and 11βHSD2, however, 11βHSD2's control is limited to the fluxes and metabolites of the 11-oxygenated pathway.

6.5 Increased expression of enzymes

Any variations in the expression levels of the enzymes can be simulated with the model. Here we look at the changes in the steady state concentrations of all intermediates and products while the expression level of the enzymes are increased. In Figure 6.14 the enzyme expressions are increased 100%, one enzyme at a time. With increased AKR1C3 expression, the steady state concentrations of T, 5αDHT, and 3αAdiol increase about 100%, while the steady state concentrations of 11KT and 11K5αDHT increase about 50% and 20% respectively. The increased AKR1C3 activity leads to a decreased steady state concentrations in 5αDione, AST, 11KA4, 11K5αDione, and 11KAST, while it has no effect on 11OH5αDione and 11OHAST. The increased expression of 17βHSD2 has little effect on any of the intermediates and products with less than 1% change in any of the metabolite concentrations. The increase in SRD5A expression increases all the metabolite steady state concentrations by around 100%, except T, 11KA4, and 11KT, which is decreased by less than 10%. The increase in 11βHSD2 expression leads to an increase in 11KA4, 11KT, 11K5αDione, 11K5αDHT, and 11KAST of around 100%.

The steady state concentrations of 11OH5αDione and 11OHAST are reduced by between 5% and 10%, while the other metabolites are unaffected. 11K5αDHT is the metabolite most influenced by the increased expression of 3αHSD; the concentration thereof is decreased by nearly 50%. The other metabolites decreased are 3αAdiol, 11OH5αDione, and 11KAST. The effect on the steady state concentration of the remaining metabolites are less than 1%. The only case where the up-regulation of an enzyme has any significant effect on lowering the synthesis of potent androgens is the up-regulation of 3αHSD which results in a decrease in 11K5αDHT synthesis of nearly 50%.

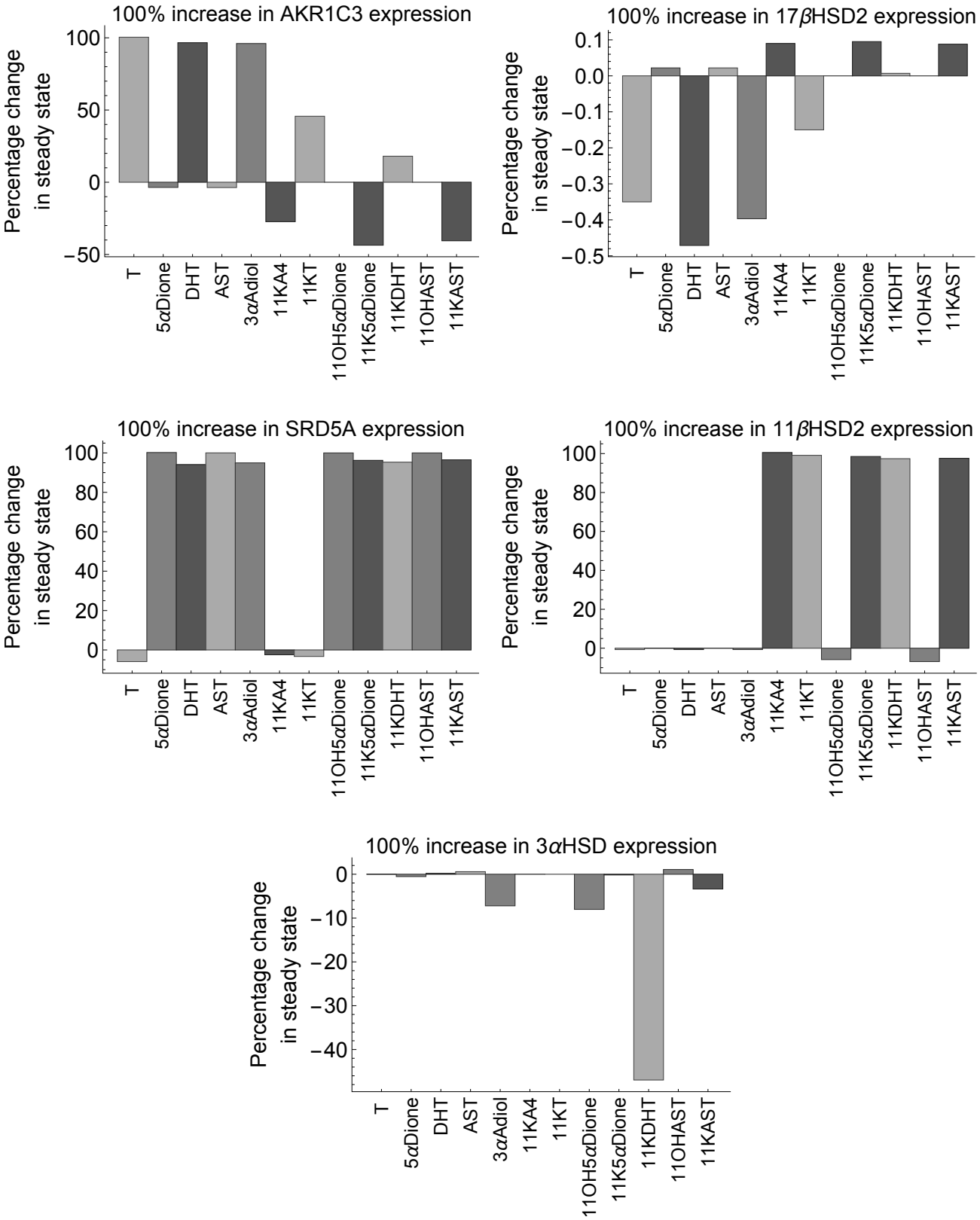


Figure 6.14: **Percentage change in steady state concentrations of metabolites with a 100% increase in enzyme expression.**

6.6 Enzyme inhibition

In Figures 6.15 and 6.16 one can see the steady state concentrations of the steroid hormones with inhibition of the enzymes AKR1C3, 17βHSD2, SRD5A, 11βHSD2, and 3αHSD. Full non-competitive inhibition of the enzymes have been simulated. Only the inhibition of AKR1C3 limits the synthesis of all potent androgens (T, 5αDHT, 11KT, and 11K5αDHT) completely. Inhibition of any of the other enzyme still leads to the synthesis of two or more of these androgens. Inhibition of 17βHSD2 has little effect on limiting the synthesis of the potent androgens. Inhibition of either SRD5A or 3αHSD limits 5αDHT and 11K5αDHT synthesis, but has little effect on the synthesis of T and 11KT. Even though SRD5A has strong control over the flux of most of the branches in the two pathways, it does not have much control over the branches leading to the synthesis of T and 11KT (Fig. 6.12). If the goal is to limit the synthesis of only 11KT and 11K5αDHT, the enzyme 11βHSD2 would be a possible target, as inhibition thereof limits the synthesis of potent 11-oxygenated androgens. However, the inhibition of 11βHSD2 has little effect on the synthesis of either T or 5αDHT.

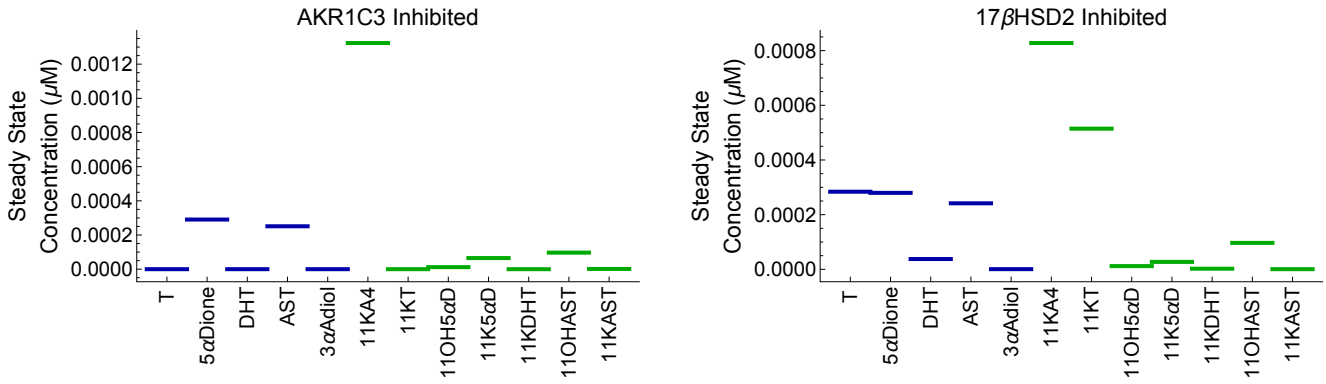


Figure 6.15: **Steady state concentrations of metabolites with 100% inhibition of AKR1C3 and 17βHSD2**

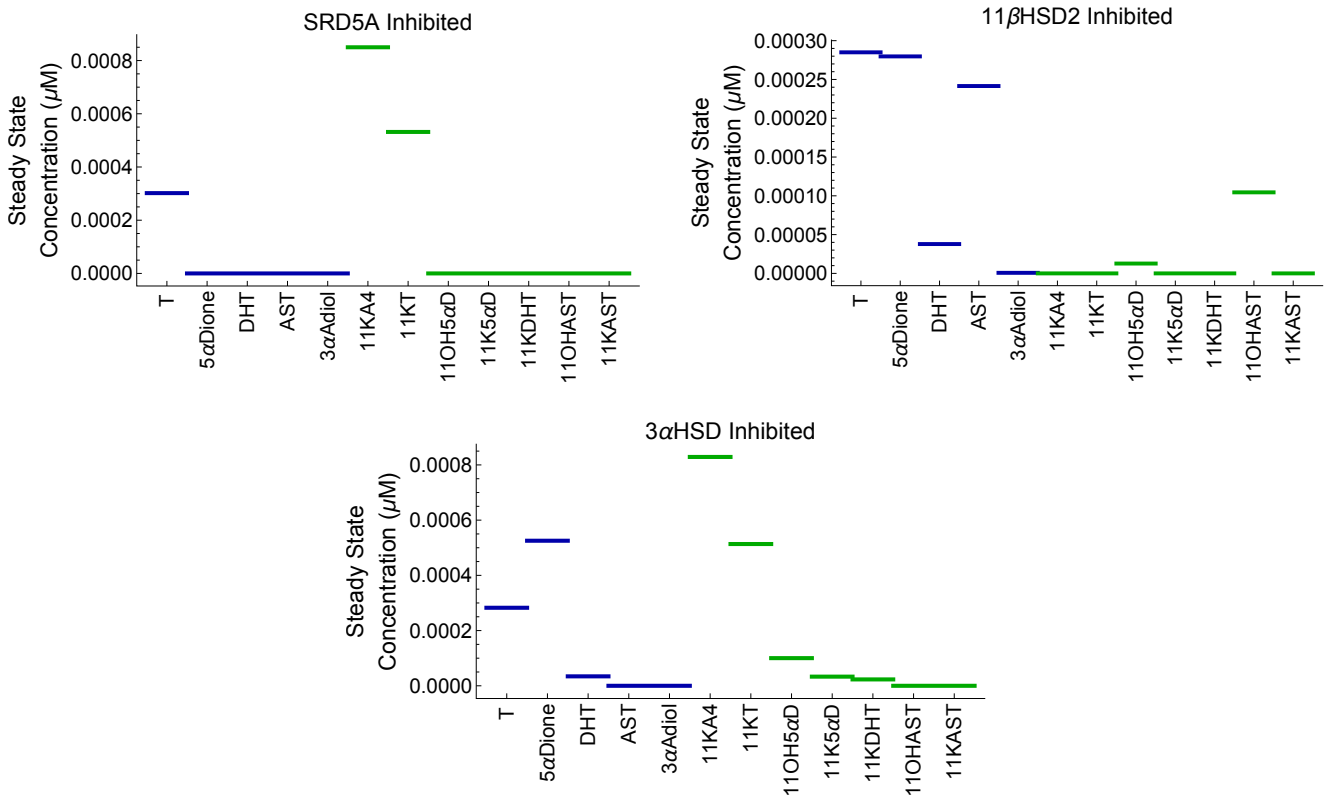


Figure 6.16: **Steady state concentrations of metabolites with 100% inhibition of SRD5A, 11βHSD2, and 3αHSD.**

Conclusion

With relatively equal serum concentrations of A4 and 11OHA4, as is seen physiologically [94], the 11-oxygenated androgens, 11KA4 and 11KT, have higher steady state concentrations than the classic androgens. MCA showed that 11βHSD2 has strong control over nearly the entire 11-oxygenated androgen pathway, but none over the classic androgen pathway, while SRD5A is the enzyme with most control over both both pathways. MCA also showed that 17βHSD2 and 3αHSD has little control over the system other than their own reactions, while AKR1C3 has strong control over some of the reactions of the classic androgen pathway, but has moderate control over some reactions of the 11-oxygenated pathway.

Model simulations of increased enzyme expression showed increased steady state concentrations of potent androgens (T, 5αDHT, 11KT, and 11K5αDHT) with increased expression of AKR1C3. Increased expression of 11βHSD2 had little effect on T and 5αDHT, but results in a great increase in 11KT and 11KT5αDHT steady state concentrations. Increased expression of SRD5A leads to a slight decrease in T and 11KT steady state concentrations, but increases the concentrations of 5αDHT and 11KT5αDHT. The only two enzymes which leads to decreased steady state concentrations of the potent

androgens with increased expression thereof are 17βHSD2 and 3αHSD, however the effects of 17βHSD2 is less than 1% change in the steady state concentrations. The increased 3αHSD expression reduces the concentration of 11K5αDHT by nearly 50%.

Model simulations of enzyme inhibition showed that AKR1C3 is the only enzyme that limits the synthesis of all four of the potent androgens with inhibition thereof. The inhibition of 11βHSD2 only limits the synthesis of 11-oxygenated androgens.

In this chapter a model is constructed for the conversion of classic androgens and 11-oxygenated androgens in C42B prostate cancer cells. The model was parameterised with values from literature and parameter values derived from experimental data of transfected cells. The model was then fitted to and validated against C42B cancer cell data.

This chapter illustrates, with the use of a computational model, that there are three possible treatment targets for limiting the synthesis of potent androgens in CRPC C42B cells. The first is the inhibition of AKR1C3, which limits the synthesis of all four potent androgens. The second is the inhibition of 11βHSD2, however, this will only limit the synthesis of 11KT and 11K5αDHT. The third possible target is the up-regulation of 3αHSD, which limits the synthesis of 11K5αDHT. The model has only been validated against C42B cancer cell data and possible future work could include the validation thereof against other prostate cancer cell lines.

Discussion

The aim of this study was to develop models for three different steroidogenic pathways and to use these models to address specific questions regarding these pathways. The first pathway that was studied was the combined Δ^4 and Δ^5 pathways in the South African Angora goat. It was previously suggested that the hypocortisolism seen in the Angora goats originate from this part of the steroidogenic pathway as these goats synthesise much more A4 relative to cortisol. Comparative *in vivo* studies in literature that were conducted on Angora goats and two hardy species, the Merino sheep and the Boer goat, showed that the main difference between the Angora goat and the hardy species is increased CYP17 activity. The 3βHSD and CYP17 enzymes are the only two enzymes catalysing the reactions of the Δ^4 and Δ^5 pathways [26–28, 118]. These enzymes are multi-functional and as such, studying the dynamics of these enzymes as they compete for the same steroid substrates is not trivial.

The research questions that were addressed in chapter 4 were the following: How does the increased CYP17 activity of the Angora goat affect the flux through the pathway? Which of the two enzymes, 3βHSD or CYP17, has the most control over the system and where does the most control lie? How does a stress response change the flux through the pathway and finally how does this all differ from a hardy species that does not suffer from hypocortisolism?

To answer these questions two models were created: one for the Merino sheep, which were used as a healthy control model and a second model which represents the altered adrenal function in the South African Angora goat. The ovine model was constructed, parameterised and validated with *in vitro* transfection data. The identifiability of the parameter values were tested, whereafter simulations of the ovine model, which mimicked the results found in [81] were used as model validation. The ovine model was also validated against animal data collected from homogenised adrenal microsomes of Merino sheep [28]. Data from the same study, but collected from Angora goats were then used to construct the Angora model. This Angora model was then validated against transfection data of another independent study [105].

After both models had been constructed and validated, the models were compared to

each other and analysed under steady state conditions. The model simulations showed that the steady state concentrations of the sex steroid precursor, A4, were higher and the cortisol precursor, 17-OHPROG, were lower in the Angora model than the ovine model. The models showed an increase in flux towards A4 and a decrease in flux towards 17-OHPROG as the activity of CYP17 is incrementally increased from that what we see in the ovine model until it finally reaches the activity level of the Angora model. The results of MCA of the two models showed that the control that both the CYP17 and the 3βHSD enzyme have on the pathway differ between the ovine and Angora models. There are two points where the change in control is most significant: the point where PREG is converted to either 17-OHPREG or PROG, and where 17-OHPREG is converted to either DHEA or 17-OHPROG by CYP17 and 3βHSD respectively. Between the ovine and Angora models there is a larger change in the flux control that both enzymes have on the Δ^5 than the reactions towards the Δ^4 part of the pathway.

The models were also analysed under steady state conditions which simulate a response to physiological stress. An increase in PREG concentration was simulated which represent the physiological response these animals have to cold stress. The relative conversion of PREG to either 17-OHPROG or A4 were analysed with increasing concentrations of PREG. The ovine model showed a physiological response one would expect to see in mammals with healthy steroidogenic function: an increase in 17-OHPROG synthesis relative to A4 synthesis with an increase in PREG. The Angora model also responded to the simulated physiological stress, however, the levels of 17-OHPROG synthesised relative to A4 is much lower than that seen in the ovine model. These results correspond to the results seen in *in vivo* studies, the South African Angora goat does respond to physiological stress by increasing the synthesis of cortisol, this response is just much lower than the response seen in hardy ovine species [26–28]. The models also show that the Δ^5 pathway is preferred in the Angora model with the flux through the Δ^5 branches being much higher than in the ovine model. Simulating a stress response also showed that the two points in the pathway mentioned above (the fork at PREG and the fork at 17-OHPREG) differ significantly between the ovine and the Angora models. These two points both contribute to maintaining the higher flux through the Δ^5 pathway seen in the Angora goat.

These results would suggest that the hypocortisolism seen in the South African Angora goat is not caused by a glitch at only one specific point in the steroidogenic pathway. These are not promising results in terms of identifying possible treatment targets. With South African mohair being of superior quality and the South African Angora goats being the only population suffering from hypocortisolism [12, 26], there might be a correlation between mohair quality and either decreased cortisol synthesis or increased sex steroid synthesis or both. The success of the breeding program aiming to increase

the hardiness of the South African Angora goat population remains to be seen [107].

The second steroidogenic pathway that was studied with the use of computational modelling is the combined pathways of the classic androgens and the 11-oxygenated androgens in the human liver. 11KT is a potent androgen and the effects of 11-oxygenated androgens should be considered along with the effects of the classic androgens as these steroids are all pooled and catalysed by the same enzymes [87, 108]. However, the 11-oxygentated androgens have not yet been studied as extensively as the classic androgens. As the 5α reduction products of both the classic androgens and the 11-oxygenated androgens are also potent androgens while the 5β reduction products are not, knowing the relative synthesis of each of these groups of steroids are of importance. The classic androgens, T and A4, are converted to their 5α reduction products, 5αDHT and 5αDione, by either SRD5A1or SRD5A2, while the 11-oxygenated androgens, 11KT and 11KA4, are converted to their 5α reduction products, 11K5αDHT and 11K5αDione. AKR1D1 converts these androgens to their 5β reduction products; T, A4, 11KA4, and 11KT are converted to 5βDHT, 5βDione, 11K5βDHT, and 11K5βDione respectively. The research question that was addressed in chapter 5 is whether the relative conversion of the classic and the 11-oxygenated androgens to their respective 5α or 5β reduction products could be predicted in the human liver with a computational model.

Chapter 5 contains the construction and validation of this model with transfection data. The model was constructed and validated with data collected from two independent transfections. The identifiability of the parameter values were also tested. The model was studied at physiological steady state with 1:1:1 enzyme expression levels. The physiological expression levels of the enzymes AKR1D1, SRD5A1, and SRD5A2 were then incorporated into the model whereafter the model was again studied at steady state.

The results showed that at enzyme expression levels of 1:1:1, the 5α reduction of T, A4, 11KT, and 11KA4 are the preferred reactions. This is because SRD5A2 has the highest efficiency towards these androgens. However, once the physiological expression levels of the enzymes are included into the model, the 5β reduction of the androgens are the preferred reactions as the expression level of AKR1D1 is nearly 16 times higher than the expression of SRD5A1 and nearly eight times higher than the expression of SRD5A2. Given physiological plasma levels if T and 11KT in men and women, the model was able to predict the conversions of these androgens to their 5α and 5β reduction products. The 5β reduction products, 5βDHT and 11K5βDHT, have the highest steady state concentrations. Of the potent 5α reduction androgen products, 5αDHT and 11K5αDHT, the levels of 5αDHT are higher in men, while women tend to have higher levels of the 11-oxygenated androgen product, 11K5αDHT.

The values used for the physiological expression of the enzymes were the relative values of the RNA expression of these enzymes in human liver cells taken from The Human Protein Atlas database. Future work would be to include the expression levels of the enzymes determined experimentally and completing additional validation experiments as the RNA expression of the enzymes are not necessarily an absolute indication of the enzyme expression levels.

The third and final steroidogenic pathway studied with the use of computational modelling is the combined dynamics of the 5αDione androgenic pathway and the 11-oxygenated androgenic pathway in C42B CRPC cells. The involvement of 11-oxygentated androgens in CRPC has increasingly gained interest in the recent years [7, 87, 108]. The research question that was addressed in chapter 6 is whether it is possible to construct a model that describe the combined dynamics of the 5αDione androgenic pathway and the 11-oxygenated androgenic pathway in C42B CRPC cells and in the process identify possible treatment targets.

The reactions of the 5αDione and the 11-oxygenated androgenic pathway is catalysed by the following enzymes: AKR1C3, 17βHSD2, SRD5A, 3αHSD, and 11βHSD2. All five enzymes are present in the 11-oxygenated androgen pathway, while only the first four are present in the 5αDione pathway. The model was constructed and parameterised with a combination of transfection data, C42B cancer cell data and previously determined parameter values. The kinetic parameters of 11βHSD2 were fitted for with the use of transfection data as these parameters had not yet been determined. The mass action kinetic parameters for 3αHSD were fitted for with the use of time course data collected from C42B cancer cells. The parameter values of SRD5A have previously been determined by our group, although they have not yet been published, and the parameter values for AKR1C3 and 17βHSD2 were taken from literature [7]. The identifiability of the 11βHSD2 parameter values were tested whereafter the model was validated against C42B cancer cell time course data.

The model was then studied at physiological steady state and MCA was completed. The results showed that the enzymes AKR1C3 and 11βHSD2 have strong control over a large part of the classic and 5αDione, and the 11-oxygenated pathways respectively, while SRD5A has strong control over a large part of both pathways. MCA showed that neither 17βHSD2 nor 3αHSD had any significant control over the reaction of the pathway other than their own reactions. The effects of increased enzyme expression as well as enzyme inhibition were simulated. Increasing the expression of these enzymes identified one possible treatment target. The up-regulation of 3αHSD reduces the concentration of the potent androgen 11K5αDHT. The inhibition of enzymes identified two possible treatment targets: the inhibition of AKR1C3 to limit the synthesis of all four potent

androgens and the inhibition of 11βHSD2 to limit the synthesis of the 11-oxygenated potent androgens, 11KT and 11K5αDHT.

Possible future work would entail parameterising the model for other CRPC cell lines and completing additional validation experiments as the enzyme kinetics might be different in other cell lines.

To complete the analysis shown in this thesis the add-on Mathematica package `IdentifiabilityAnalysis` was created. The aim of creating the package was to allow the user to parameterise an ODE based model with time course type data with very little additional coding required. In the process the user can also test the identifiability of the model parameters and create profile likelihood plots which give a visual representation of the parameter identifiability and confidence intervals.

The package consists of five functions. The function `objectiveFuncion` returns the SSR value between the model and data. This function returns a single value as it calculates the SSR for a specific set of parameter values. The function `minimizeObjectiveFunction` is similar to `objectiveFuncion`, however, this function scans over the possible solution space while minimising the SSR values and in the process returns the optimal parameter value estimations. The profile likelihood for all the parameters are calculated by the function `profileLikelihood` and the `likelihoodRatioTest` function completed the likelihood ratio test on the profile likelihood values at a specific Chi-square distribution value, as selected by the user, corresponding to one degree of freedom and the desired confidence level. The function `likelihoodRatioPlot` conducts the same test as the function `likelihoodRatioTest`, however, the results are returned as a profile likelihood plot.

This study confirms the importance of each step of the model construction process. Chapter 4 highlighted the crucial role that identifiability analysis plays in experimental design. In the other two research chapters in this study, identifiability analysis was only used to test the identifiability of newly determined parameters, as well as calculating their respective confidence intervals. For the construction of the ovine model in chapter 4, identifiability analysis aided in showing where there was a lack of experimental data. Identifiability analysis allowed us to determine which parameters were not identifiable and which datasets would need to be added to improve the identifiability. Therefore only a few, very specific additional experimental datasets were needed. Chapter 5 is an example of where the inclusion of a few non-identifiable parameters in a model still results in a usable model. Although the ovine model in chapter 4 also included non-identifiable parameters, these parameters corresponded with a branch in the steroidogenic pathway with the lowest steroid flux and which is often omitted during

the modelling of that particular pathway. The non-identifiable parameters seen in chapter 5 played a more prominent role in the model. It must be said that all cases of non-identifiability were practical non-identifiability, which only resulted in very wide parameter confidence intervals. As such, the use of these parameters in the model, although technically not ideal, still resulted in a useful model. Chapter 6 shows the difficulties in combining parameter values from literature with newly determined parameter values from experimental data. This chapter highlights the importance of sufficient validation data and the pit fall of using construction and validation data that are too similar.

There are similarities in the structure of the androgenic pathways discussed in this study, however, there is variation in the roles of these pathways in different organs and organisms. These differences can be attributed to varied protein expression levels and differences in the levels of steroids in circulation. Therefore the comparison of models could be extended to the steroidogenic pathways in different organs of different organisms. The results in this study reinforce the importance of statistical and analytical techniques used during the computational modelling of steroid hormone biosynthesis.

Appendix A

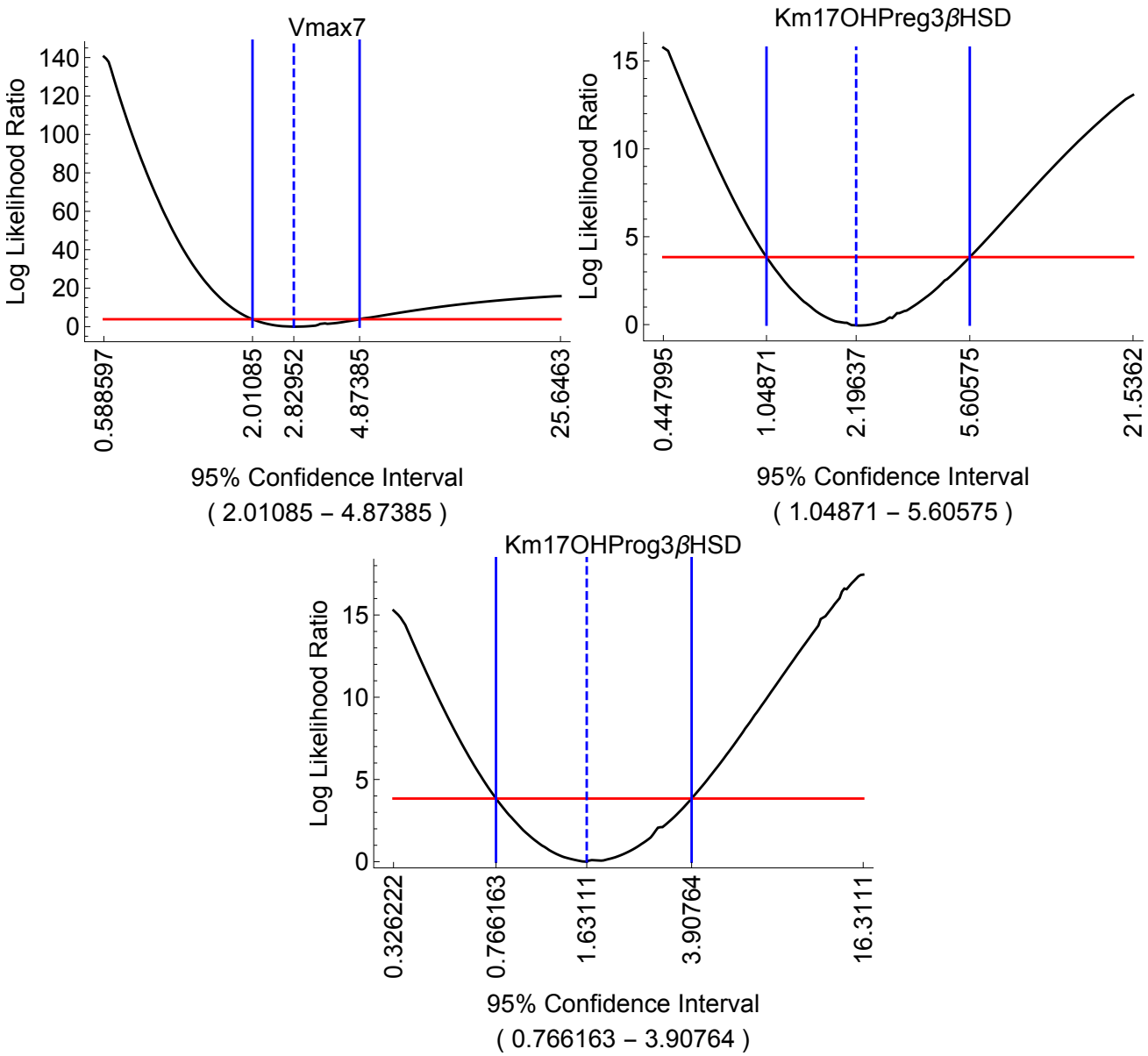


Figure 6.17: Identifiability analysis of 3βHSD $Vmax_2$, $Km_{17OHPreg}$ and $Km_{17OPHProg}$ for the conversion of 17OH-PREG to 17OH-PROG. The red lines indicate the 95% confidence cut-off value.

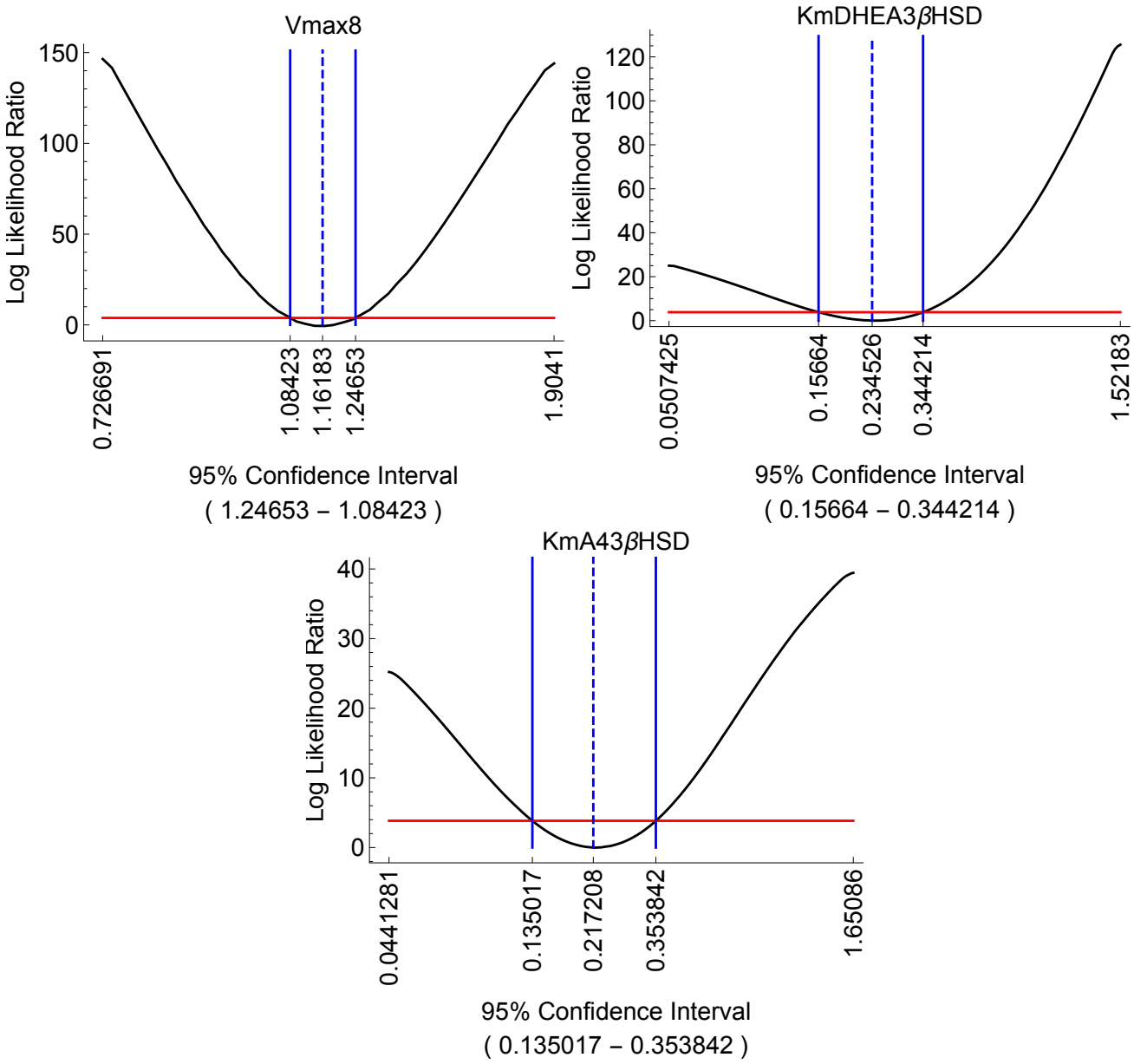


Figure 6.18: Identifiability analysis of 3βHSD $Vmax_3$, Km_{DHEA} and Km_{A4} for the conversion of DHEA to A4. The red lines indicate the 95% confidence cut-off value.

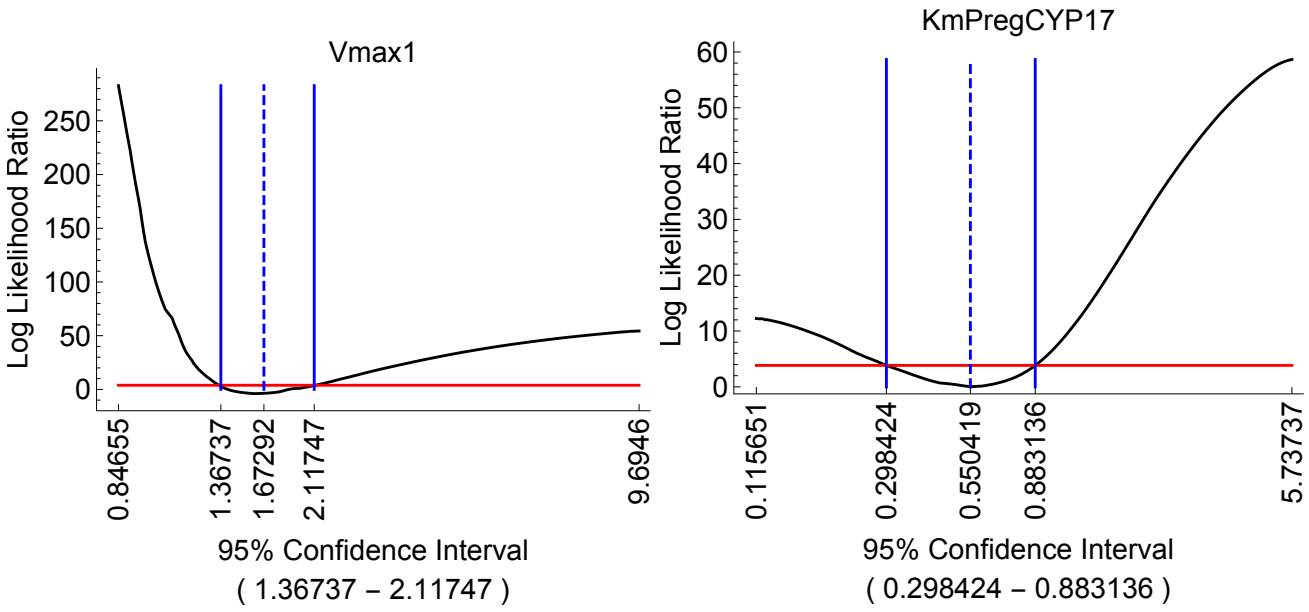


Figure 6.19: Identifiability analysis of CYP17 $Vmax_1$ and Km_{Preg} for the conversion of PREG to 17OH-PREG. The red lines indicate the 95% confidence cut-off value.

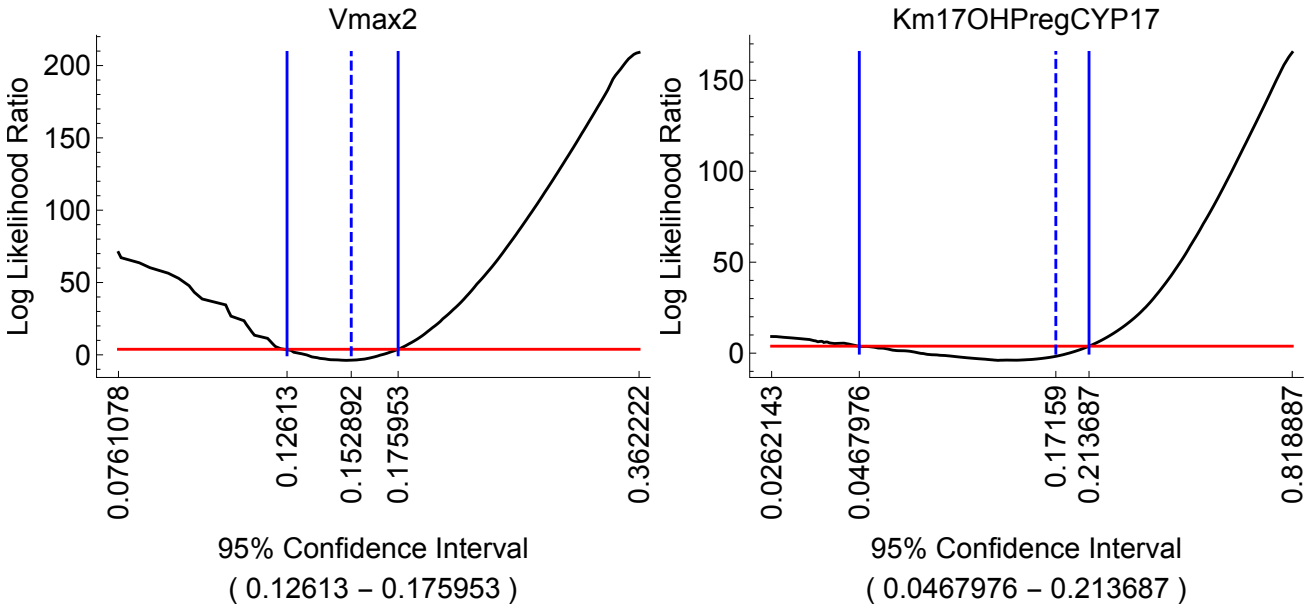


Figure 6.20: Identifiability analysis of CYP17 $Vmax_2$ and $Km_{17OHPreg}$ for the conversion of 17OH-PREG to DHEA. The red lines indicate the 95% confidence cut-off value.

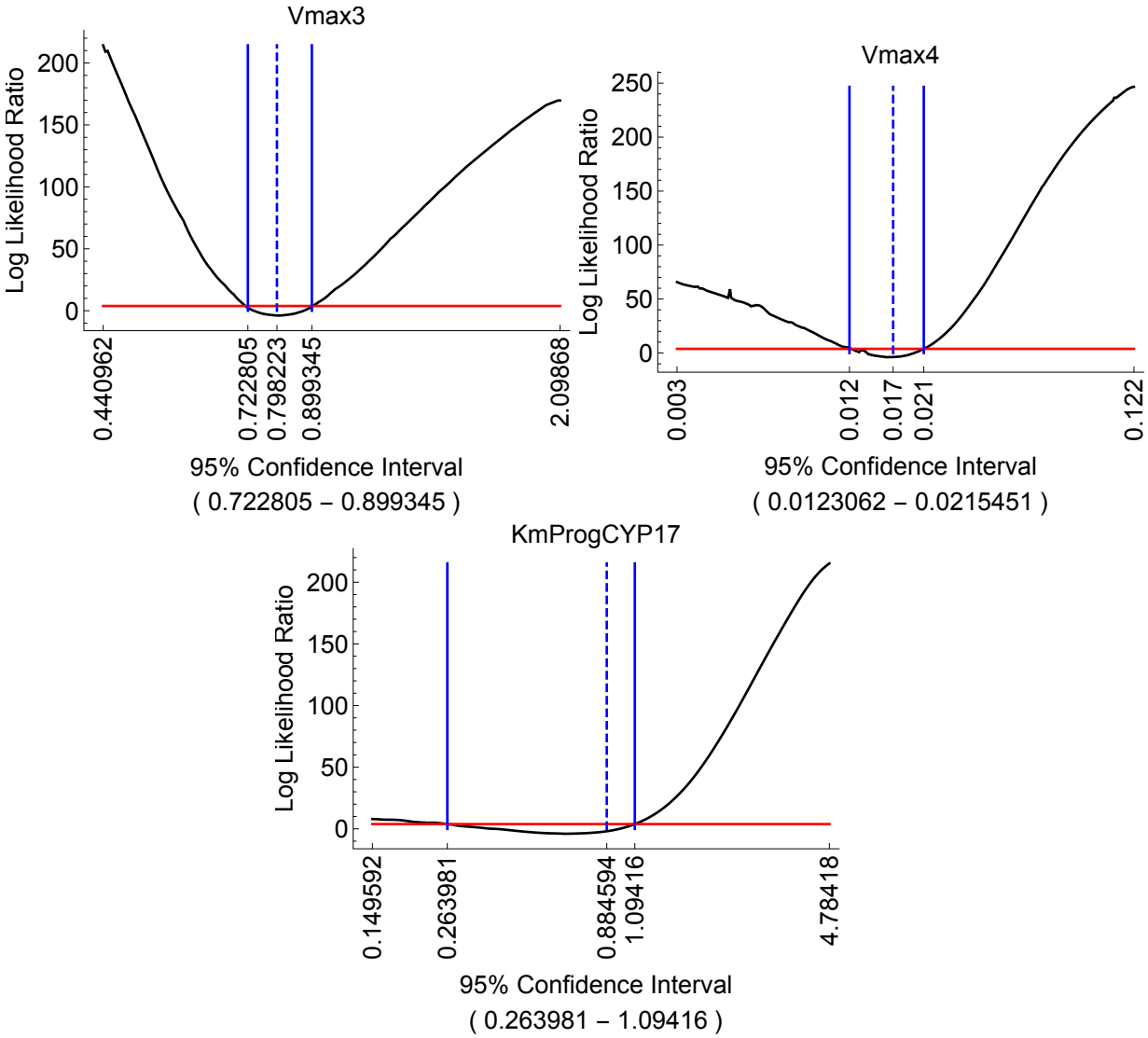


Figure 6.21: Identifiability analysis of CYP17 $Vmax_3$, $Vmax_4$ and Km_{Prog} for the conversion of PROG to 17OH-PROG and 16OH-PROG. The red lines indicate the 95% confidence cut-off value.

Appendix B

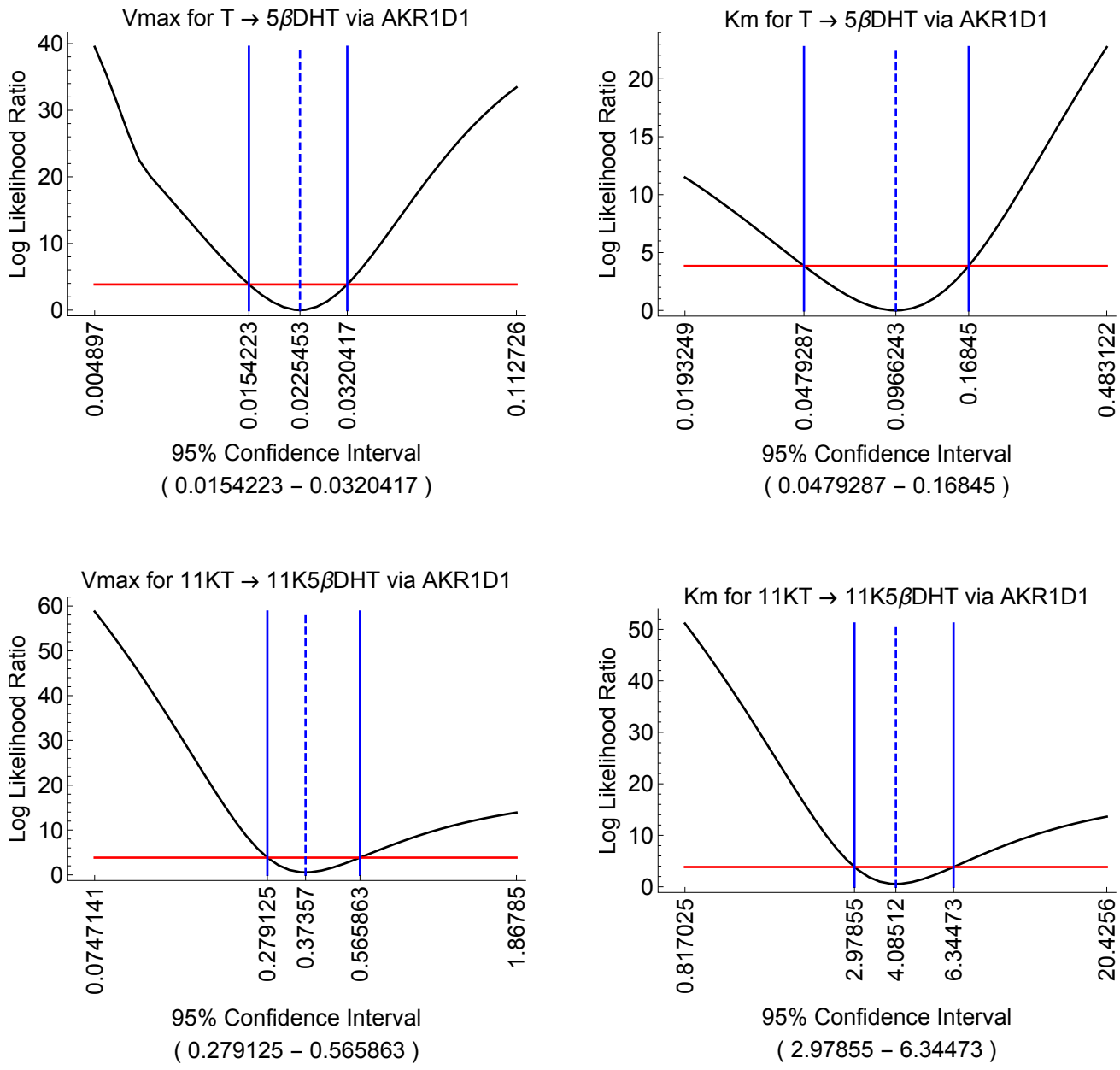


Figure 6.22: Identifiability analysis of AKR1D1 $Vmax$ and Km for the conversion of T to 5βDHT and the conversion of 11KT to 11K5βDHT. The red lines indicate the 95% confidence cut-off value.

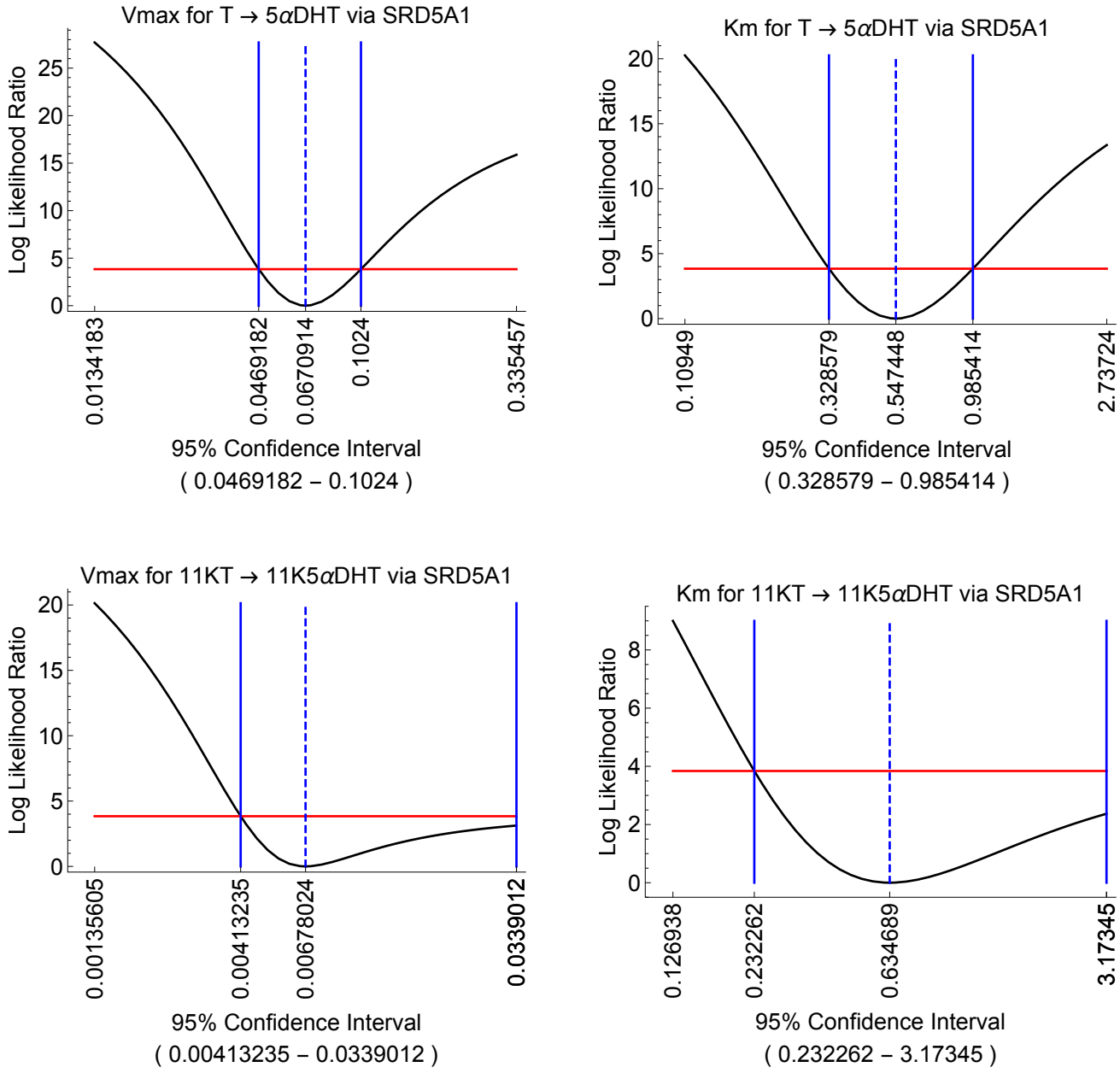


Figure 6.23: Identifiability analysis of SRD5A1 $Vmax$ and Km for the conversion of T to 5αDHT and the conversion of 11KT to 11K5αDHT. The red lines indicate the 95% confidence cut-off value.

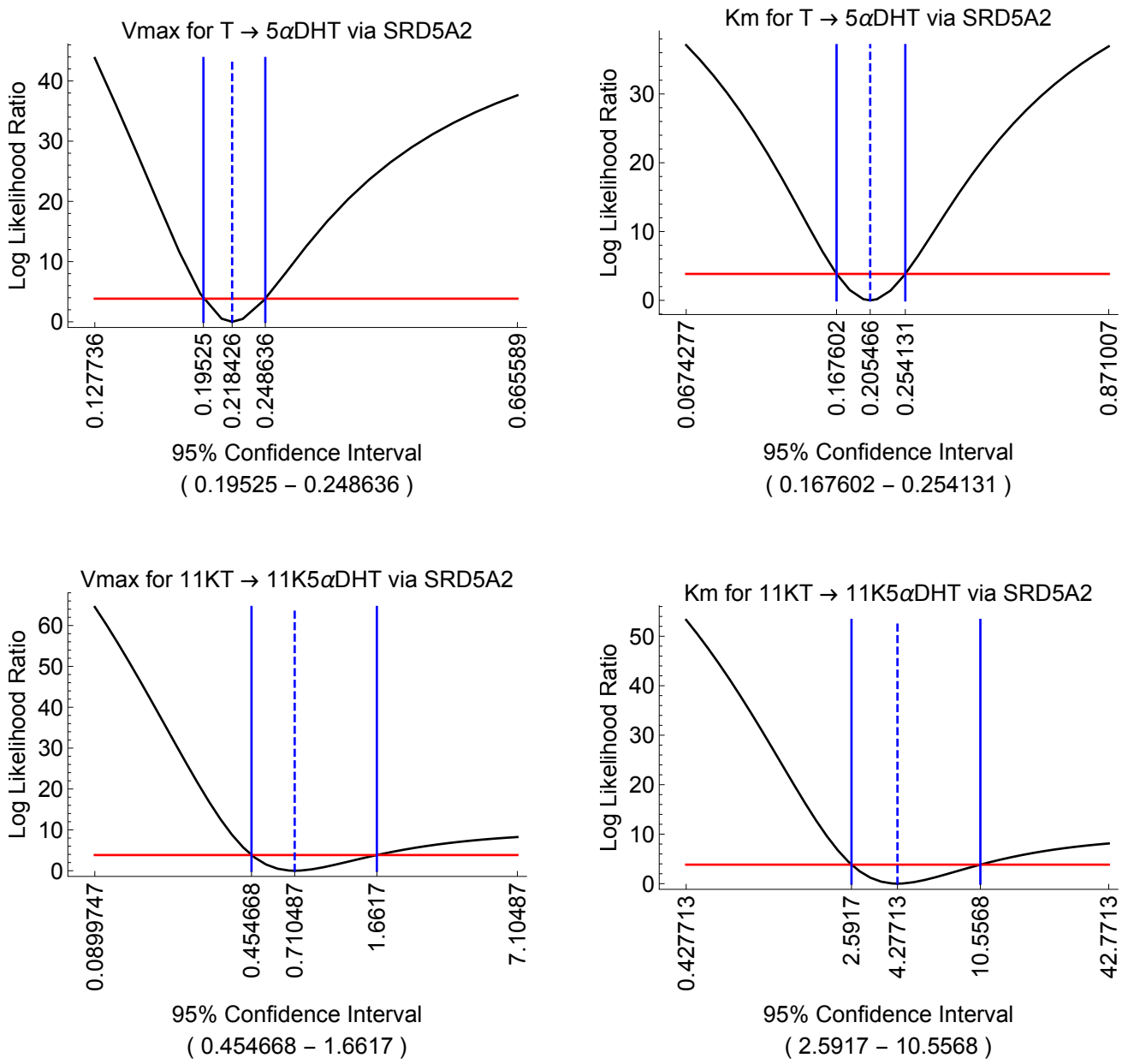


Figure 6.24: Identifiability analysis of SRD5A2 $Vmax$ and Km for the conversion of T to 5αDHT and the conversion of 11KT to 11K5αDHT. The red lines indicate the 95% confidence cut-off value.

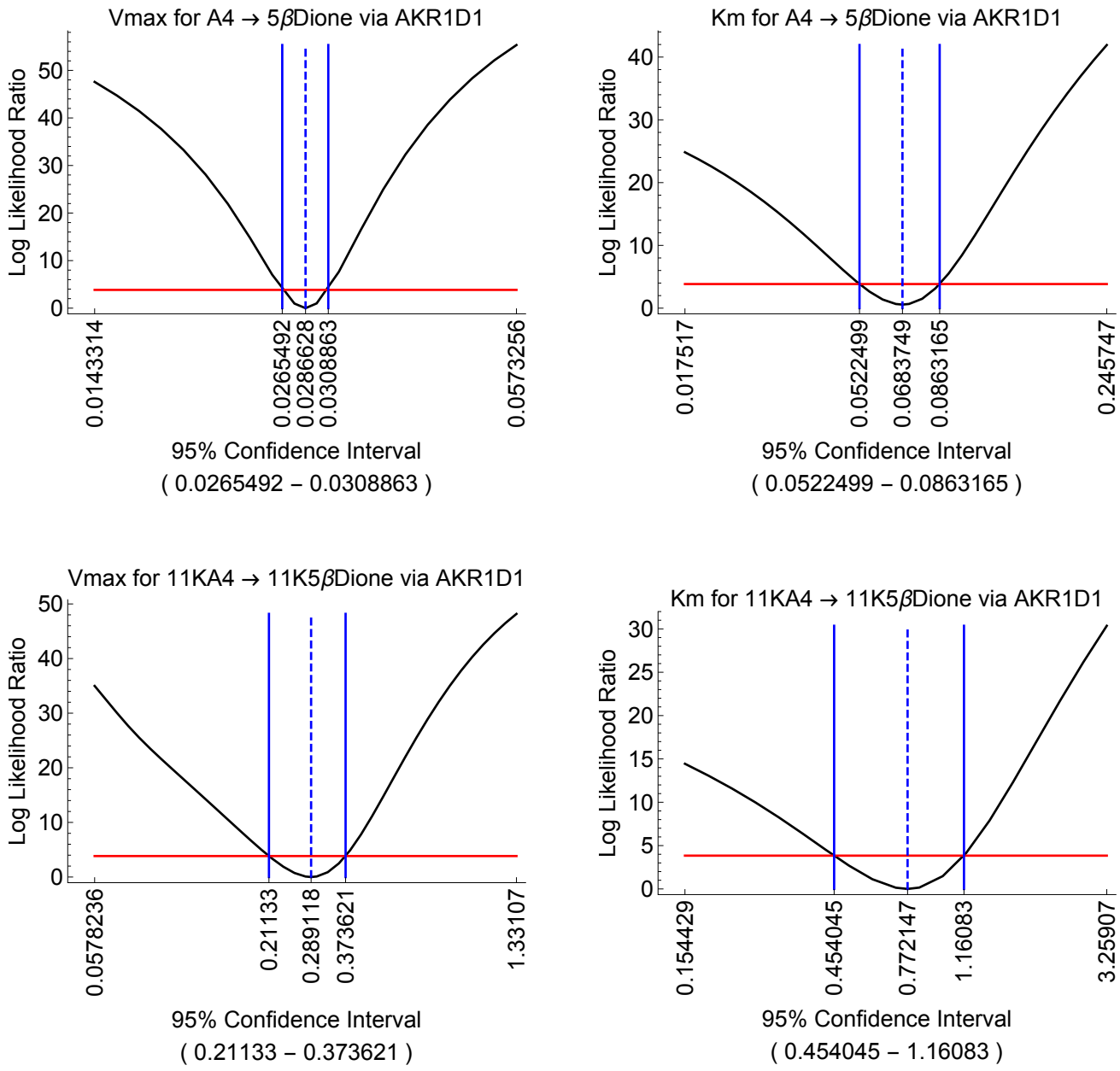


Figure 6.25: Identifiability analysis of AKR1D1 $Vmax$ and Km for the conversion of A4 to 5βDione and the conversion of 11KA4 to 11K5βDione. The red lines indicate the 95% confidence cut-off value.

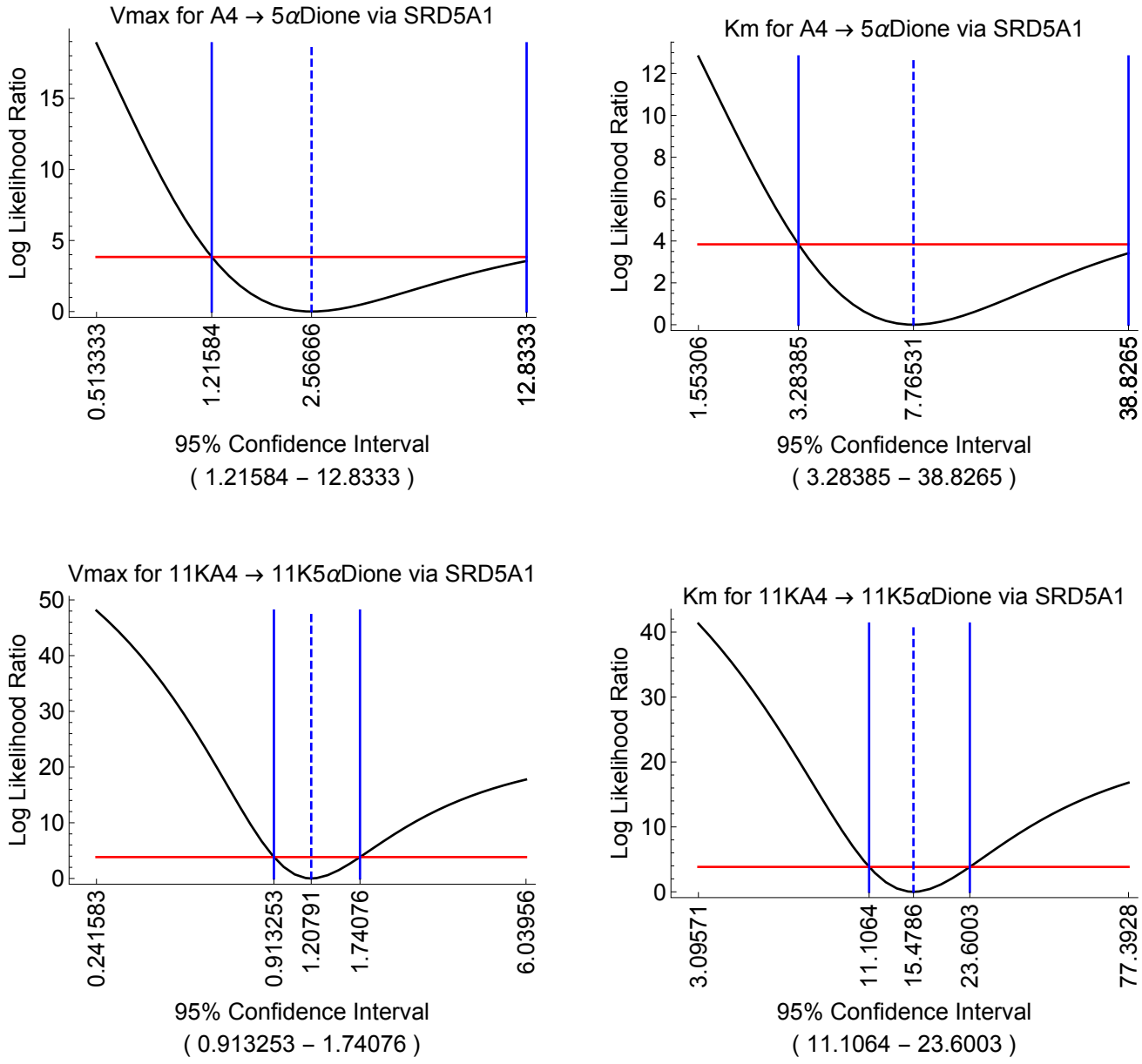


Figure 6.26: Identifiability analysis of SRD5A1 $Vmax$ and Km for the conversion of A4 to 5αDione and the conversion of 11KA4 to 11K5αDione. The red lines indicate the 95% confidence cut-off value.

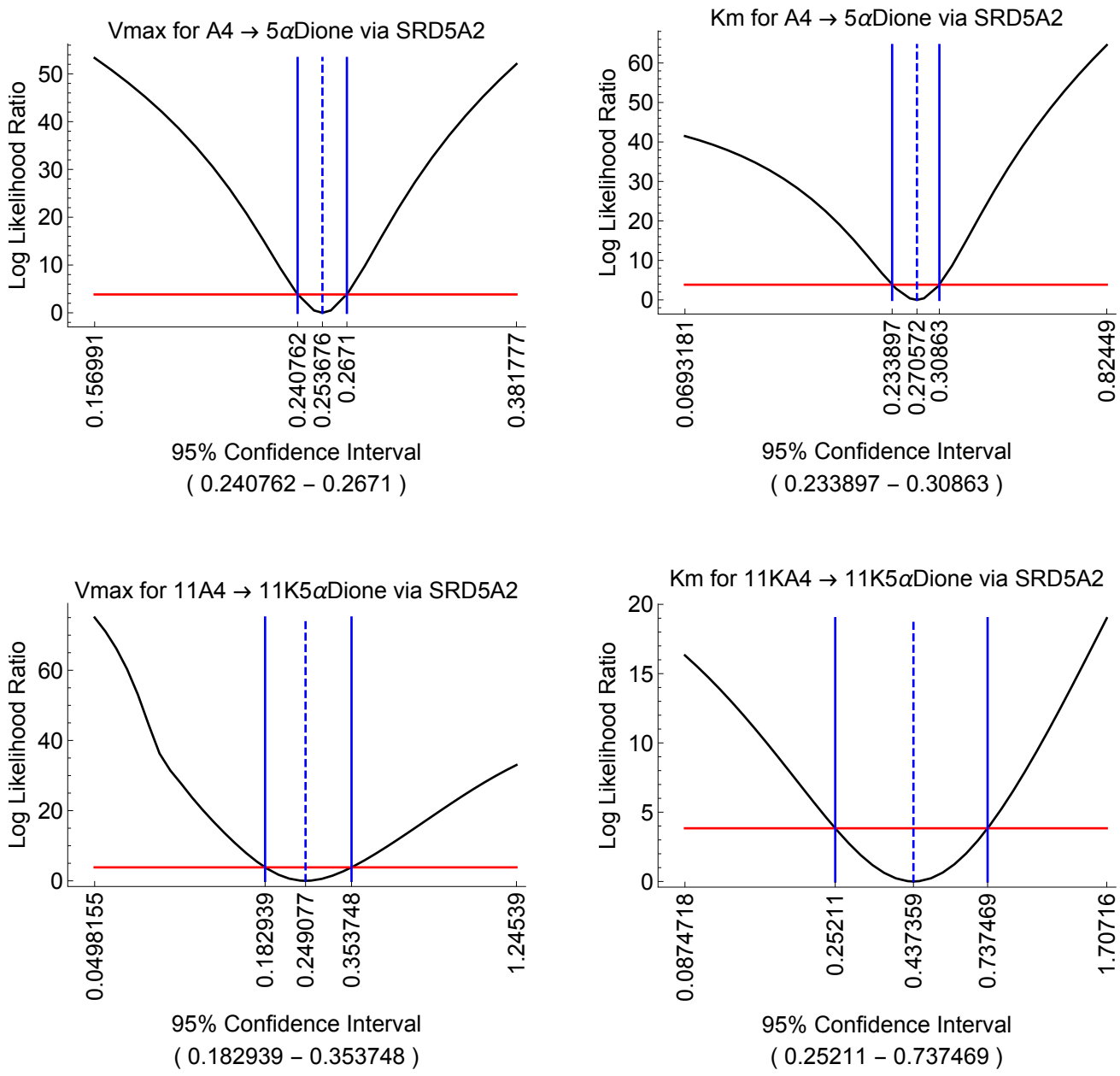


Figure 6.27: Identifiability analysis of SRD5A2 $Vmax$ and Km for the conversion of A4 to 5αDione and the conversion of 11KA4 to 11K5αDione. The red lines indicate the 95% confidence cut-off value.

Bibliography

[1] Africander, D., Storbeck, K., 2018, "Steroid Metabolism in Breast Cancer: Where Are We and What Are We Missing?" *Mol Cell Endocrinol.* **466**, 86-97.

[2] Alper, J. S., Gelb, R. I., 1990, "Standard Errors and Confidence Intervals in Nonlinear Regression: Comparison of Monte Carlo and Parametric Statistics" *J Phys Chem.* **94**, 4747-4751.

[3] Alper, J. S., Gelb, R. I., 1991, "Monte Carlo Method for the Determination of Confidence Intervals: Analysis of Nonnormally Distributed Errors in Sequential Experiments" *J Phys Chem.* **95**, 108-111.

[4] Altman, D., Machin, D., Bryant, T., Gardner, M., 2000, *Statistics with Confidence: Confidence Intervals and Statistical Guidelines*, Wiley, New York.

[5] Ashyraliyev, M., Jaeger, J., Blom, J. G., 2008, "Parameter Estimation and Determinability Analysis Applied to *Drosophila* Gap Gene Circuits" *BMC Syst Biol.* **2**(1), 83-101.

[6] Auchus, R. J., Lee, T. C., Miller, W. L., 1998, "Cytochrome b_5 Augments the 17,20-Lyase Activity of Human P450c17 Without Direct Electron Transfer" *J Biol Chem.* **273**(6), 3158-3165.

[7] Barnard, M., Quanson, J. L., Mostaghel, E., Pretorius, E., Snoep, J. L., Storbeck, K., 2018, "11-Oxygenated Androgen Precursors are the Preferred Substrates for Aldo-Keto Reductase 1C3 (AKR1C3): Implications for Castration Resistant Prostate Cancer" *J Steroid Biochem.* **183**, 192-201.

[8] Bates, D. M., Watts, D. G., 1988, *Nonlinear Regression Analysis and Its Applications*, 2nd. Wiley, New York.

[9] Becker, S., Chubb, C., Ewing, L., 1980, "Mathematical Model of Steroidogenesis in Rat and Rabbit Testes" *Am J Physiol-Reg I.* **239**(1), 184-195.

[10] Bird, I. M., Conley, A. J., 2002, Steroid Biosynthesis: Enzymology, Integration and Control. In: Mason, J. I., ed. *Genetics of Steroid Biosynthesis and Function.* London: Taylor and Francis, Chapter 1: 1-35.

[11] Bloem, L. M., Storbeck, K., Schloms, L., Swart, A. C., 2013, "11β-Hydroxyandrostenedione Returns to the Steroid Arena: Biosynthesis, Metabolism and Function" *Molecules*. **18**(11), 13228-13244.

[12] Braun, A. L., 1999, 'Cashmere anf Mohair Quality and Value-Adding Potential by SMME's', *Commercialization of South African Goats, Witwatersrand Agriculture Show*, 7 April 1999.

[13] Breen, M., Breen, M. S., Terasaki, N., Yamazaki, M., Lloyd, A. L., Conolly, R. B., 2011, "Mechanistic Computational Model of Steroidogenesis in H295R Cells: Role of Oxysterols and Cell Proliferation to Improve Predictability of Biochemical Response to Endocrine Active Chemicals - Metyrapone" *Toxicol Sci*. **123**(1), 80-93.

[14] Breen, M. S., Breen, M., Terasaki, N., Yamazaki, M., Conolly, R. B., 2010, "Computational Model of Steroidogenesis in Human H295R Cells to Predict Biochemical Response to Endocrine-Active Chemicals: Model Development for Metyrapone" *Environ Health Persp*. **118**(2), 265-272.

[15] Breen, M. S., Villeneuve, D. L., Breen, M., Ankley, G. T., Conolly, R. B., 2007, "Mechanistic Computational Model of Ovarian Steroidogenesis to Predict Biochemical Responses to Endocrine Active Compounds" *Ann Biomed Eng*. **35**(6), 970-981.

[16] Briggs, G. E., Haldane, J. B., 1925, "A Note on the Kinetics of Enzyme Action" *Biochem J*. **19**, 338-339.

[17] Brubaker, T. A., Tracy, R., Pomernacki, C.L., 1978, "Linear Parameter Estimation" *Anal Chem*. **50**(11), 1017-1024.

[18] Brubaker, T. A., O'Keefe, K.R., 1979, "Nonlinear Parameter Estimation" *Anal Chem*. **51**(13), 1385-1388.

[19] Cobelli, C., Finkelstein, L., Carson, E. R., 1982, "Mathematical Modelling of Endocrine and Metabolic Systems: Model Formulation, Identification and Validation" *Math Comput Simulat*. **24**, 442-451.

[20] Cole, S. R., Chu, H., Greenland, S., 2014, "Maximum Likelihood, Profile Likelihood, and Penalized Likelihood: A Primer" *Am J Epidemiol*. **179**(2), 252-260.

[21] Conley, A. J., Bird, I. M., 1997, "The Role of Cytochrome P450 17α-Hydroxylase and 3β-Hydroxysteroid Dehydrogenase in the Integration of Gonadal and Adrenal Steroidogenesis via the Δ5 and Δ4 Pathways of Steroidogenesis in Mammals" *Biol Reprod*. **56**, 789-799.

[22] Cook, S. F., Fiandalo, M. V., Watt, D. S., Wu, Y., Mohler, J. L., Bies, R. R., 2018, "Mathematical Modeling of Intracrine Androgen Metabolism in Prostate Cancer: Methodological Aspects" *The Prostate*. **78**(14), 1069-1076.

[23] Dorfman, D. D., Alf, E., 1969, "Maximum-Likelihood Estimation of Parameters of Signal-Detection Theory and Determination of Confidence Intervals - Rating-Method Data" *J Math Psychol.* **6**, 487-496.

[24] Duffy, D. M., Hess, D. L., Stouffer, R. L., 1994, "Acute Administration of a 3β-Hydroxysteroid-Dehydrogenase Inhibitor to Rhesus Monkeys at the Midluteal Phase of the Menstrual Cycle: Evidences for Possible Autocrine Regulation of Primate Corpus Luteum by Progesterone" *J Clin Endocrinol Metab.* **79**, 1587-1597.

[25] Duggleby, R. G., 1995, "Analysis of enzyme progress curves by non-linear regression" *Methods Enzymol.* **249**, 61-90.

[26] Engelbrecht, Y., 1999, 'A Biochemical Investigation of Hypoadrenocorticism in the South African Angora Goat', PhD Thesis, University of Stellenbosch.

[27] Engelbrecht, Y., Herselman, T., Louw, A., Swart, P., 2000, "Investigation of the Primary Cause of Hypoadrenocorticism in South African Angora Goats (*Capra aegagrus*): A Comparison with Boer Goats (*Capra Hircus*) and Merino Sheep (*Ovis Aries*)" *J Anim Sci.* **78**, 371-379.

[28] Engelbrecht, Y., Swart, P., 2000, "Adrenal Function in Angora Goats: A Comparative Study of Adrenal Steroidogenesis in Angora Goats, Boer Goats, and Merino Sheep" *J Anim Sci.* **78**, 1036-1046.

[29] Ezertas, A. A., 2009, 'Sensitivity Analysis Using Finite Difference and Analytical Jacobians', MSc Thesis, Graduate School of Natural and Applied Science, Middle East Technical University.

[30] Faller, D., Klingmüller, U., Timmer, J., 2003, "Simulation Methods for Optimal Experimental Design in Systems Biology" *Simulation.* **79**(12), 717-725.

[31] Fevold, H. R., 1967, "Regulation of the Adrenal Cortex Secretory Pattern by Adrenocorticotropin" *Science.* **156**(3783), 1753-1755.

[32] Fevold, H. R., Lorence, M. C., McCarthy, J. L., Trant, J. M., Kagimoto, M., Waterman M. R., Mason, J. I., 1989, "Rat $P450_{17\alpha}$ from Testis: Characterization of a Full-Length cDNA Encoding a Unique Steroid Hydroxylase Capable of Catalyzing Both Δ^4- and Δ^5-Steroid-17,20-Lyase Reactions" *Mol Endocrinol.* **3**(6), 968-975.

[33] Friedman, J., Hastie, T., Tibshirani, R., 2001, *The Elements of Statistical Learning*, Vol. 1, No. 10, New York: Springer series in statistics.

[34] Fröhlich F., Theis F. J., Hasenauer J., 2014, "Uncertainty Analysis for Non-identifiable Dynamical Systems: Profile Likelihoods, Bootstrapping and More", in P. Mendes, J. O. Dada, K. Smallbone (eds.), *Computational Methods in Systems Biology. CMSB 2014. Lecture Notes in Computer Science*, Springer, Cham.

[35] Gardner, M. J., Altmen, D. G., 1986, "Confidence Intervals Rather Than P Values: Estimation Rather Than Hypothesis Testing" *Brit Med J.* **292**, 746-750.

[36] Gavin, H. P., 2013, The Levenberg-Marquardt Method for Nonlinear Least Squares Curve-fitting Problems, Duke University, Durham on 9 October 2013.

[37] Gesztelyi, R., Zsuga, J., Kemeny-Beke, A., Varga, B, Juhasz, B., Tosaki, A., 2012, "The Hill Equation and the Origin of Quantitative Pharmacology" *Arch Hist Exact Sci.* **66**, 427-438.

[38] Goosen, P., 2012, 'The Influence of 3βHSD on Adrenal Steroidogenesis and the Factors which Influence its Activity', PhD Thesis, University of Stellenbosch.

[39] Goosen, P., Swart, A. C., Storbeck, K., Swart, P., 2010, "Hypocortisolism in the South African Angora Goat: The Role of 3βHSD" *Mol Cell Endocrinol.* **315**, 182-187.

[40] Heinrich, R., Rapoport, T. A., 1974, "A Linear Steady-State Treatment of Enzymatic Chains: General Properties, Control and Effector Strength" *Eur J Biochem.* **42**, 89-95.

[41] Hennig, P., Kiefel, M., 2013, "Quasi-Newton Methods: A New Direction" *J Mach Learn Res.* **14**, 843-865.

[42] Henri, V., 1903, "Lois générales de l'action des diastases", Hermann, Paris.

[43] Herselman, M. J., Pieterse, D. M. E., 1992, "Effek van Kortisolbehandeling op Angorabokke" In *Proc SA Soc Anim Prod.* **31**, South Africa.

[44] Hofmeyr, J. H. S., 2001, "Metabolic Control Analysis in a Nutshell", In *Proceedings of the 2nd International conference on systems biology*, Madison, Wisconsin: Omnipress. November, pp. 291-300.

[45] Holzbeierlein, J., Lal, P., LaTulippe, E., Smith, A., Satagopan, J., Zhang, L., Ryan, C., Smith, S., Scher, H., Scardino, P., Reuter, V., Gerald, W. L., 2004, "Gene Expression Analysis of Human Prostate Carcinoma During Hormonal Therapy Identifies Androgen-Responsive Genes and Mechanisms of Therapy Resistance" *Am J Pathol.* **164**, 217-227.

[46] Hornberg, J. J., Bruggeman, F. J., Bakker, B. M., Westerhoff, H. V., 2007, Metabolic Control Analysis to Identify Optimal Drug Targets, In: Boshoff, H. I, Barry, C., E. eds. *Progress in Drug Research: Systems Biology Approaches in Infectious Diseases: Volume 64*. Basel: Springer Science and Business Media, Chapter 7: 171-190.

[47] Hough, D., Cloete, S. W. P., Storbeck, K., Swart, A. C., Swart, P., 2013, "Cortisol Production in Sheep is Influenced by the Functional Expression of Two Cytochrome

P450 17α-Hydroxylase/17,20-Lyase (CYP17) Isoforms" *J Anim Sci.* **91**(3), 1193-1206.

[48] Hough, D., Storbeck, K., Cloete, S. W. P., Swart, A. C., Swart, P., 2015, "Relative Contribution of P450c17 Towards the Acute Cortisol Response: Lessons from Sheep and Goats" *Mol Cell Endocrinol.* **408**, 107-113.

[49] Hovland, P., Norris, B., Roh, L., Smith, B., 1999, "Developing a Derivative-enhanced Object-orientated Toolkit for Scientific Computations", In M. E. Henderson, C. R. Anderson and S. L. Lyons (eds.), *Proceedings of the SIAM Workshop on Object Orientated Methods for Inter-operable Scientific and Engineering Computing*, Yorktown Heights, New York, October 21-23, 1998, pp. 129-137.

[50] Huelsenbeck, J. P., Crandall, K. A., 1997, "Phylogeny Estimation and Hypothesis Testing Using Maximum Likelihood" *Annu Rev Ecol Syst.* **28**, 437-466.

[51] Huet, S., Bouvier, A., Poursat, M. A., Jolivet, E., 2006. *Statistical tools for nonlinear regression: a practical guide with S-PLUS and R examples*, Springer Science & Business Media.

[52] Imamichi, Y., Yuhki, K., Orisaka, M., Kitano, T., Mukai, K., Ushikubi, F., Taniguchi, T., Umezawa, A., Miyamoto, K., Yawaza, T., 2016, "11-Ketotestosterone is a Major Androgen Produced in Human Gonads" *J Clin Endocrinol Metab.* **101**, 3582-3591.

[53] Ioslovich, I., Moran, M. I. R., Gutman, P., 2010, "Identification of a Nonlinear Dynamic Biological Model Using the Dominant Parameter Selection Method" *J Frankl Inst.* **347**, 1001-1014.

[54] Iott, J., Haftka, T., Adelman, H. M., 1985, "Selecting Step Sizes in Sensitivity Analysis by Finite Differences", Technical Memorandum, NASA Langley Research Center, August 1985, pp. 1-11.

[55] Jacquez, J. A., Greif, P., 1985, "Numerical Parameter Identifiability and Estimability: Integrating Identifiability, Estimability, and Optimal Sampling Design" *Math Biosci.* **77**, 201-227.

[56] Johnson, K. A., 2013, "A century of enzyme kinetic analysis, 1913 to 2013" *FEBS Lett.* **587**, 2753-2766.

[57] Kacser, H., Burns, J. A., 1973, "The Control of Flux" *Symp Soc Exp Biol.* **27**, 65-104.

[58] Keesman, K. J., 2011, *System identification: an introduction*, Springer, New York.

[59] Kozlov, K., Surkova, S., Myasnikova, E., Reinitz, J., Samsonova, M., 2012, "Modeling of Gap Gene Expression in *Drosophila Kruppel* Mutants" *PLOS Comput Biol.* **8**(8), 1-17.

[60] Krekels, E., H., J., Rower, J. E., Constance, J. E., Knibbe, C. A. J., Sherwin, C. M. T., 2017, "Chapter 8 - Hepatic Drug Metabolism in Pediatric Patients", in Xie, W. (ed.), *Drug Metabolism in Diseases*, Academic Press, 181-206.

[61] Kreutz, C., Raue, A., Kaschek, D., Timmer, J., 2013, "Profile Likelihood in Systems Biology" *FEBS J.* **280**(11), 2564-2571.

[62] Kyriakides, E., Heydt, G. T., 2006, "Calculating Confidence Intervals in Parameter Estimation: A Case Study" *IEEE T Power Deliver.* **21**(1), 508-509.

[63] Lee-Robichaud, P., Wright, J. N., Akhtar, M. E., Akhtar, M., 1995, "Modulation of the Activity of Human 17α-Hydroxylase-17,20-Lyase (CYP17) by Cytochrome b_5: Endocrinological and Mechanistic Implications" *Biochem J.* **308**, 901-908.

[64] Li, P., Vu, Q. D., 2013, "Identification of Parameter Correlations for Parameter Estimation in Dynamic Biological Models" *BMC Syst Biol.* **7**(91), 1-12.

[65] Mason, J. I., 1993, "Cytochrome P450 as a Target for Therapeutic Inhibitors" *Biochem Soc T.* **21**, 1057-1060.

[66] Mathur, R., 2012, 'An Analytical Approach to Computing Step Sizes for Finite-Difference Derivatives', PhD Thesis, University of Texas at Austin.

[67] Meeker, W. Q., Escobar, L. A., 1995, "Teaching About Approximate Confidence Regions Based on Maximum Likelihood Estimation" *Am Stat.* **49**(1), 48-53.

[68] Michaelis, L., Menten, M.L., 1913, "Die kinetik der invertinwirkung" *Biochem Z.* **49**, 333-369.

[69] Miller, W. L., 2002, "Androgen Biosynthesis form Cholesterol to DHEA" *Mol Cell Endocrinol.* **198**, 7-14.

[70] Miller, W. L., Auchus, R. J., 2011, "The Molecular Biology, Biochemistry, and Physiology of Human Steroidogenesis and Its Disorders" *Endocr Rev.* **32**(1), 81-151.

[71] Mohler, J. L., Gregory, C. W., Ford, O. H., Kim, D., Weaver, C. M., Petrusz, P., Wilson ,E. M., French, F. S., 2004, "The Androgen Axis in Recurrent Prostate Cancer" *Clin Cancer Res.* **10**, 440-448.

[72] Moisan, A. M., Ricketts, M. L., Tardy, V., Desrochers, M., Mébarki, F., Chaussain, J., Cabrol, S., Raux-Demay, M. C., Forest, M. G., Sippell, W. G., Peter, M., Morel, Y., Simard, J., 1999, "New Insight into the Molecular Basis of 3β-Hydroxysteroid Dehydrogenase Deficiency: Identification of Eight Mutations in the *HSD3B2* Gene in Eleven Patients from Seven New Families and Comparison of the Functional Properties of Twenty-Five Mutant Enzymes" *J Clin Endocr Metab.* **84**(12), 4410-4425.

[73] Montgomery, D. C., Peck, E. A., 1992, *Introduction to Linear Regression Analysis*, Wiley, New York.

[74] Mpyana, M., 2017, "Market Profile of the South African Mohair Industry" In *Trade Probe* **71**, Department of Agriculture, Forestry and Fisheries (November 2017).

[75] Mukaka, M. M., 2012, "Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research" *Malawi Med J.* **23**(3), 69-71.

[76] Murphy, C. A., Rose, K. A., Thomas, P., 2005, "Modeling Vitellogenesis in Female Fish Exposed to Environmental Stressors: Predicting the Effects of Endocrine Disturbance due to Exposure to a PCB Mixture and Cadmium" *Reprod Toxicol.* **19**, 395-409.

[77] Nanba, A. T., Rege, J., Ren, J., Auchus, R. J. Rainey, W. E., Turcu, A. F., 2019, "11-Oxygenated C19 Steroids Do Not Decline with Age in Women" *J Clin Endrocr Metab.* **104**(7), 2615-2622.

[78] Nelder, J. A., Mead, R., 1965, "A Simplex Method for Function Minimization" *Comput J.* **7**(4), 308-313.

[79] Nguyen, P. T. T., 2012, 'Mathematical Modelling of Steroid Synthesis in Steroid-Producing Tissues', PhD Thesis, University of Auckland.

[80] Nguyen, P. T. T., Conley, A. J., Sneyd, J., Lee, R. S. F., Soboleva, T. K., Shorten, P. R., 2013, "The Role of Enzyme Compartmentalization on the Regulation of Steroid Synthesis" *J Theor Biol.* **332**, 52-64.

[81] Nguyen, P. T. T., Lee, R. S. F., Conley, A. J., Sneyd, J., Soboleva, T. K., 2012, "Variation in 3β-Hydroxysteroid Dehydrogenase Activity and in Pregnenolone Supply Rate Can Paradoxically Alter Androstenedione Synthesis" *J Steroid Biochem.* **128**, 12-20.

[82] Nijhout, H. F., Reed, M. C., Ulrich, C. M., 2008, "Mathematical Models of Folate-Mediated One-Carbon Metabolism" *Vitam Horm.* **79**, 45-82.

[83] Pan, J., Fang, K., 2002, *Growth Curve Models and Statistical Diagnostics*, Springer, New York.

[84] Pang, S., 2001, "Congenital Adrenal Hyperplasia Owing to 3β-Hydroxysteroid Dehydrogenase Deficiency" *Endocrin Metab Clin.* **30**, 81-99.

[85] Patocs, A., Liko, I., Varga, I., Gergics, P., Boros, A., Futo, L., Kun, I., Bertalan, R., Toth, S., Pazmany, T., Toth, M., Szücs, N., Horanyi, J., Glaz, E, Racz, K., 2005, "Novel Mutation of the CYP17 Gene in Two Unrelated Patients with Combined 17α-Hydroxylase/17,20-lyase Deficiency: Demonstration of Absent Enzyme Activity by

Expressing the Mutnt CYP17 Gene and by Three-Dimensional Modeling" *J Steroid Biochem Molec Biol.* **97**, 257-265.

[86] Payne, A. H., Hales, D. B., 2004, "Overview of Steroidogenic Enzymes in the Pathway from Cholesterol to Active Steroid Hormones" *Endocr Rev.* **25**(6), 947-970.

[87] Pretorius, E., Arlt, W., Storbeck, K., 2017, "A New Dawn for Androgens: Novel Lessons from 11-Oxygenated C19 Steroids" *Mol Cell Endocrinol.* **441**, 76-85.

[88] Rateitschak, K., Winter, F., Lange, F., Jaster, R., Wolkenhauer, O., 2012, "Parameter Identifiability and Sensitivity Analysis Predict Targets for Enhancements of STAT1 Activity in Pancreatic Cancer and Stellate Cells" *PLOS Comput Biol.* **8**(12), 1-14.

[89] Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., 2009, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood" *Bioinformatics.* **25**(15), 1923-1929.

[90] Rosencrantz, M. A., Coffler, M. S., Haggan, A., Duke, K. B., Donohue, M. C., Shayya, R. F., Su, H. I., Chang, R. J., 2011, "Clinical Evidence for Predominance of Delta-5 Steroid Production in Women with Polycystic Ovary Syndrome" *J Clin Endocrinol Metab.* **96**(4), 1106-1113.

[91] Rosol, T. J., Gröne, A., 2016, Endocrine Glands, In: Maxie, G. ed. *Jubb, Kennedy & Palmer's Pathology of Domestic Animals: Volume 3.* Missouri: Elsevier, Chapter 3: 269-357.

[92] Ruckstuhl, A., 2010, *Introduction to Nonlinear Regression: Study guide*, Institut für Datenanalyse und Prozessdesign, Zürcher Hochschule für Angewandte Wissenschaften, Winterthur.

[93] Saito, R., Terasaki, N., Yamazaki, M., Masutomi, N., Tsutsui, N., Okamoto, M., 2016, "Estimation of the Mechanism of Adrenal Action of Endocrine-Disrupting Compounds Using a Computational Model of Adrenal Steroidogenesis in NCI-H295R Cells" *J Toxicol.* **2016**, 1-19.

[94] Schiffer, L., Arlt, W., Storbeck, K., 2018, "Intracrine Androgen Biosynthesis, Metabolism and Action Revisited" *Mol Cell Endocrinol.* **465**, 4-26.

[95] Segel, I.H., Segel, A.H., 1976, *Biochemical calculations: how to solve mathematical problems in general biochemistry*, Wiley, New York.

[96] Selgrade, J. F., Schlosser, P. M., 1999, "A Model for the Production of Ovarian Hormones During the Menstrual Cycle" *Fields Inst Commun.* **21**, 429-446.

[97] Selgrade, J. F., Schlosser, P. M., 2000, "A Model of Gonadotropin Regulation During the Menstrual Cycle in Women: Qualitative Features" *Environ Health Persp.* **108**(5), 873-881.

[98] Shafi, A. A., Yen, A. E., Weigel, N. L., 2013, "Androgen Receptors in Hormone-Dependent and Castration-Resistant Prostate Cancer" *Pharmacol Ther.* **140**, 223-238.

[99] Simard, J., Rheaume, E., Mebarki, F., Sanchez, R., New, M. I., Morel, Y., Labrie, F., 1995, "Molecular Basis of Human 3β-Hydroxysteroid Dehydrogenase Deficiency" *J Steroid Biochem Molec Biol.* **52**(1-6), 127-138.

[100] Snoep, J. L., Mendes, P., Westerhoff, H., 1999, "Teaching Metabolic Control Analysis and Kinetic Modelling" *The Biochemist.* **25**, 25-28.

[101] Soleymani, F., Sharifi, M., 2009, "A New Derivative-Free Quasi-Secant Algorithm for Solving Non-Linear Equations" *Int J Math Comp Sci.* **3**(7), 554-556.

[102] Steward, P. M., Newell-Price, J. D. C., 2003, The Adrenal Cortex. In Larsen, P. R., Kronenberg, H. M., Melmed, S., Polonsky, K. S., eds. *Williams Textbook of Endocrinology.* Philadelphia: Saunders, Chapter 15: 490-555.

[103] Storbeck, K., Swart, A. C., Slabbert, J. T., Swart, P., 2007, "The Identification of Two CYP17 Alleles in the South African Angora Goat" *Drug Metab Rev.* **39**, 467-480.

[104] Storbeck, K., 2008, 'The Influence of Dual CYP17 Expression on Adrenal Steroidogenesis in the South African Angora Goat', PhD Thesis, University of Stellenbosch.

[105] Storbeck, K., Kolar, N. W., Stander, M., Swart, A. C., Prevoo, D., Swart, P., 2008, "The Development of an Ultra Performance Liquid Chromatography-Coupled Atmospheric Pressure Chemical Ionization Mass Spectrometry Assay for Seven Adrenal Steroids" *Anal Biochem.* **372**, 11-20.

[106] Storbeck, K., Swart, A. C., Snyman, M. A., Swart, P., 2008, "Two CYP17 genes in the South African Angora goat (*Capra hircus*) - The Identification of Three Genotypes that Differ in Copy Number and Steroidogenic Output" *FEBS J.* **275**, 3934-3943.

[107] Storbeck, K., Swart, A. C., Swart, P., 2009, "CYP17 Causes Hypocortisolism in the South African Angora Goat" *Mol Cell Endocrinol.* **300**, 121-125.

[108] Storbeck, K., Bloem, L. M., Africander, D., Schloms, L., Swart, P., Swart, A.C., 2013, "11β-Hydroxydihydrotestosterone and 11-Ketodihydrotestosterone, Novel C19 Steroids with Androgenic Activity: A Putative Role in Castration Resistant Prostate Cancer?" *Mol Cell Endocrinol.* **377**, 135-146.

[109] Stryhn, H., Christensen, J., 2003, "Confidence Intervals by the Profile Likelihood Method, with Applications in Veterinary Epidemiology" In *Proceedings of the 10th International Symposium on Veterinary Epidemiology and Economics*, Vina del Mar, 208.

[110] Swart, A. C., Storbeck, K., Swart, P., 2010, "A Single Amino Acid Residue, Ala 105, Confers 16α-Hydroxylase Activity to Human Cytochrome P450 17α-Hydroxylase/17,20-Lyase" *J Steroid Biochem Mol Biol.* **119**, 112-120.

[111] Swart, P., Lombard, N., Swart, A. C., Van der Merwe, T., Murry, B. A., Nicol, M., Mason, J. I., 2003, "Ovine Steroid 17α-hydroxylase Cytochrome P450: Characteristics of the Hydroxylase and Lyase Activities of the Adrenal Cortex Enzyme" *Arch Biochem Biophys.* **409**, 145-152.

[112] Symes, E. K., Milroy, E. J. G., Mainwaring, W. I. P., 1978, "The Nuclear Uptake of Androgen by Human Benign Prostate In Vitro: Action of Antiandrogens" *J Urol.* **120**, 180-183.

[113] The Human Protein Atlas 2019, *AKR1D1*, The Human Protein Atlas, viewed 16 October 2019, <https://www.proteinatlas.org/ENSG00000122787-AKR1D1/tissue/liver>

[114] The Human Protein Atlas 2019, *SRD5A1*, The Human Protein Atlas, viewed 16 October 2019, <https://www.proteinatlas.org/ENSG00000145545-SRD5A1/tissue/liver>

[115] The Human Protein Atlas 2019, *SRD5A2*, The Human Protein Atlas, viewed 16 October 2019, <https://www.proteinatlas.org/ENSG00000277893-SRD5A2/tissue/liver>

[116] Turcu, A., Smith, J. M., Auchus, R., Rainey, W. E., 2014, "Adrenal Androgens and Androgen Precursors: Definition, Synthesis, Regulation and Physiological Actions" *Compr Physiol.* **4**(4), 1369-1381.

[117] Uusipaikka, E., 2008, *Confidence Intervals in Generalized Regression Models*, CRC Press, Boca Raton.

[118] Van Rensburg, S. J., 1971, "Reproductive Physiology and Endocrinology of Normal and Habitually Aborting Angora Goats" *Onderstepoort J Vet Res.* **38**(1), 1-62.

[119] Van Schalkwyk, F. J., 2016, 'Construction and Analysis of a Kinetic Model of the Ovine Adrenal Steroidogenic Subsystem Comprising of CYP17 and 3β-HSD', MSc Thesis, University of Stellenbosch.

[120] Venzon, D. J., Moolgavkar, S. H., 1988, "A Method for Computing Profile-Likelihood-Based Confidence Intervals" *Appl Statist.* **37**(1), 87-94.

[121] Weber, P., Hasenauer, J., Allgöwer, F., Radde, N., 2011, "Parameter Estimation and Identifiability of Biological Networks Using Relative Data" *IFAC Proceedings Volumes*. **44**(1), 11648-11653.

[122] Weems, Y. S., Bridges, P. J., Lea Master, B. R., Vincent, D. L., Weems, C. W., 1999, "Effect of Trilostane on Prostagladins (PG) E and $F_{2\alpha}$ (PGE; $PGF_{2\alpha}$), Estradiol-17βand Progesterone Secretion and Pregnancy of 90 Days Ovariectomized Pregnant Ewes" *Prostagladin Other Lipid Mediators*. **58**, 77-86.

[123] Wentzel, D., Viljoen, K. S., Botha, L. J. J., 1979, "Physiological and Endocrinological Reactions to Cold Stress in the Angora Goat" *Agroanimalia*. **11**, 19-22.

[124] Wentzel, D., 1982, "Alkalinization of Grain for High Energy Feeding of Cold-Stressed Angora Goats" In *Proc 3rd Int Conf Goat Prod & Disease*, Tucson, Arizona.

[125] Westerhoff, H. V., Kolodkin, A., Conradie, R., Wilkinson, S. J., Bruggeman, F. J., Krab, K., Van Schuppen, J., Hardin, H., Bakker, B. M., Moné, M. J., Rybakova, K. N., Eijken, M., Van Leeuwen, H. J., Snoep, J. L., 2009, "Systems Biology Towards Life In Silico: Mathematics of the Control of Living Cells" *J Math Biol*. **58**(1-2), 7-34.

[126] White, A., Tolman, M., Thames, H. D., Withers, H. D., Mason, K. A., Transtrum, M. K., 2016, "The Limitations of Model-Based Experimental Design and Parameter Estimation in Sloppy Systems" *PLOS Comput Biol*. **12**(12), 1-26.

[127] Yao, K. Z., Shaw, B. M., Kou, B., McAuley, K. B., Bacon, D. W., 2003, "Modeling Ethylene/Butane Copolymerization with Multi-site Catalyst: Parameter Estimability and Experimental Design" *Polym React Eng*. **11**(3), 563-588.