Plasmodium falciparum dihydrofolate synthase-folylpolyglutamate synthase (DHFS-FPGS): Gene synthesis and recombinant expression

by

Linda Coetzee

Submitted in partial fulfilment of the requirements for the degree *Magister Scientiae*
in the Faculty of Natural and Agricultural Sciences
Department of Biochemistry
University of Pretoria
Pretoria
December 2003

Table of Contents

Acknowledgements i
Abbreviations ii
List of tables iv
List of figures v

Chapter 1

Literature Review

1.1 The burden of malaria 1
1.2 Malaria pathogenesis 2
1.3 The multifaceted problem of malaria 3
1.3.1 Vector control 3
1.3.2 Host genetic factors 4
1.3.3 Pathogen control 4
1.3.3.1 Cutting-edge technologies 4
1.3.3.2 Vaccines 5
1.3.3.3 Drugs 6
1.4 Drug targets 7
1.4.1 Haemoglobin degradation in the lysosomal food vacuole 8
1.4.2 Apicoplast metabolism 9
1.4.3 Limited electron transport due to an acrystate mitochondrion 9
1.4.4 Cytosolic and membrane bound targets 9
1.5 The folate pathway 10
1.5.1 Folate metabolic organisation 10
1.5.2 Current status of the antifolates 11
1.5.3 Drug synergy 13
1.5.4 Exogenous folate utilisation 13
1.5.5 Putative drug targets within the folate pathway 14
1.5.6 Bifunctional dihydrofolate synthase-folylpolyglutamate synthase (DHFS-FPGS) as a possible drug target 14
1.6 Research aims 17

Chapter 2

PCR-mediated synthesis of the *P. falciparum dhfs-fpgs* gene

2.1 Introduction 18
2.2 Materials and methods 21

2.2.1 Oligonucleotides and primers 21
2.2.1.1 Oligonucleotide design 22
2.2.1.2 Oligonucleotide- and primer stock solutions and nucleic acid concentration determination 22
2.2.2 PCR gene synthesis 23
2.2.3 Agarose gel electrophoresis and purification 27
2.2.4 Cloning 28
2.2.4.1 The pGEM T Easy vector system 28
2.2.4.2 The pMOS*Blue* vector system 30
2.2.5 Plasmid isolation and identification of recombinant clones 31
2.2.5.1 STET-prep plasmid isolation 31
2.2.5.2 High Pure Plasmid kit plasmid isolation 32
2.2.6 Sequencing 32
2.2.7 Construction of the full-length gene 33
2.2.8 Gene repair strategies 34
2.2.8.1 Quarter gene segments 34
2.2.8.2 Site-directed mutagenesis of point mutated PCR products 35
2.2.8.3 Restriction-ligation of correct fragments 36
2.2.8.4 Cassette mutagenesis PCR 36
2.2.8.5 Construction of the full-length gene from correct quarter fragments 38
2.3 Results 40
2.3.1 Oligonucleotide design 40
2.3.2 Optimisation of the overlap-extension assembly step 40
2.3.3 Optimisation of the number of assembly cycles 42
2.3.4 Construction of the full-length gene from half fragments 44
2.3.5 Gene repair 44
2.3.6 Construction of the full-length gene from correct quarter fragments 46
2.4 Discussion 48
2.4.1 Overlap-extension PCR as agene synthesis method 48
2.4.2 Optimisation of the assembly PCR parameters 49
2.4.3 Optimisation of the number of assembly cycles 49
2.4.4 Construction of the full-length gene 50
2.4.5 Gene repair 51
2.4.6 Construction of the full-length gene from quarter fragments 52

Chapter 3

Expression of synthetic *P. falciparum* dihydrofolate synthase-folylpolyglutamate synthase (DHFS-FPGS)

3.1 Introduction 53
3.2 Materials and methods 56
3.2.1 Constructs, vectors and cell lines 56

3.2.2 Protein expression 57
3.2.3 Determination of the protein concentration 58
3.2.4 SDS PAGE analysis 58
3.2.5 Functional complementation 59
3.2.6 Partial protein purification 60
3.2.6.1 Unfolding/refolding protocol 60
3.2.6.2 Affinity chromatography 60
3.2.6.3 Size exclusion high performance liquid chromatography 61
3.3 Results 62
3.3.1 Expression of a variety of *dhfs-fpgs* constructs 62
3.3.2 Functional complementation 64
3.3.3 Preliminary purification studies 66
3.3.3.1 Affinity purification of refolded C-terminal His_6-tagged DHFS-FPGS from inclusion bodies 66
3.3.3.2 Affinity purification of soluble C-terminal His_6-tagged DHFS-FPGS 67
3.3.3.3 Size exclusion high performance liquid chromatography 67
3.4 Discussion 70
3.4.1 Expression of various DHFS-FPGS constructs 70
3.4.2 Functional complementation 71
3.4.3 Partial purification 72
3.4.3.1 Unfolding and refolding of inclusion bodies 72
3.4.3.2 Affinity purification of C-terminal His_6-tagged DHFS-FPGS. 72
3.4.4 Future prospects 73

Chapter 4

In silico analysis of dihydrofolate synthase-folylpolyglutamate synthase (DHFS-FPGS)

4.1 Introduction 74
4.2 Methods 77
4.2.1 Sequence alignments 77
4.2.2 Structure predictions 77
4.3.Results 79
4.3.1 Inter-species DHFS and FPGS alignments and phylogenetic analysis 79
4.3.2 Conservation of DHFS-FPGS within the *Plasmodium* species and comparison with human FPGS 86
4.3.3 Secondary structure prediction of *P. falciparum* DHFS-FPGS 88
4.3.4 DHFS-FPGS hydrophobicity profile 90
4.4 Discussion 91
4.4.1 Sequence conservation of DHFS-FPGS 91
4.4.2 Predicted secondary structure 92

Chapter 5

Concluding discussion

5.1 'Discover, Develop, Deliver' 93
5.2 Antifolates:New targets, old pathway 93
5.3 *P. falciparum* DHFS-FPGS: an attractive drug target 94
5.3 The aims of this study: obtaining sufficient amounts of active *P. falciparum* DHFS-FPGS 94
5.5 A look into the near and distant future 96

Summary 98
Opsomming 100
References 102
Appendices 114

Acknowledgements

To Professor Louw, my promoter and mentor, who made this project as much his own as it was mine, shared in all the ups and downs, from whom I learnt all about science and much about life in general, who believed in my abilities and gave me opportunities far beyond that and to Prof. Neitz, my co-promoter, for his sharp mind and valuable comments. To the NRF for rewarding me a scholarship that enabled these studies.

To Professor Hyde, Dr. Tanya Aspinall and the rest of the UMIST research team, for receiving and hosting me with such warmth, for all their good advice and assistance and their willingness to cooperate and work with me towards our common goal.

To all our staff, lecturers and students at UP, especially Jaco de Ridder for the design of the oligonucleotides, who gave me a new understanding and appreciation of computers, Gordon Wells who generously donated time from his own thesis to help me in the Bioinformatics department, as well as Sandra, Ben and Christine for helping me with the technical and experimental detail and Lyn-Marie for her much valued advice and insights.

To my parents, who realised the importance of a good education and who set some of their dreams aside to provide me with this, for their unconditional love and support. To my grandparents, family and friends for their continual support, interest and understanding. To Heinrich for his undaunted support and belief in my abilities, who shared with me all my failures and successes and loved me nevertheless.

Lastly, to our Father, Who loved us enough to become one of us, Who created this wonderful universe we have yet to begin to understand, Who knows all the answers to our myriad of questions, Who grants us brief glimpses of His perfection through the subjects we study and explore…

Abbreviations

ACP	acyl carrier protein
ARMS PCR	amplification refractory mutation system PCR
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	basepair
BSA	bovine serum albumin
CPG	controlled pore glass
CS	circumsporozoite
DHFR	dihydrofolate reductase
DHFS	dihydrofolate synthase
DHPS	dihydropteroate synthase
DMT	4,4'dimethoxytrityl
DNA	deoxyribonucleic acid
DOXP	deoxyxylulose phosphate
dNTP	deoxynucleoside triphosphate
dTMP	deoxythymidine monophosphate
DTT	dithiotreitol
dUMP	deoxyuridine monophosphate
E.C.	Enzyme Commission
EDTA	(ethylenedinitrilo)tetraacetic acid
FPGS	folylpolyglutamate synthase
GTP	guanosine triphosphate
GTP-CH	GTP cyclohydrolase
HLA	human leukocyte antigen
HNN	Hierarchial Neural Networks
HPLC	high performance liquid chromatography
HRP	histidine rich protein
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside
LB	Luria-Bertani medium
LDH	lactate dehydrogenase
MMV	Medicines for Malaria Venture
MVI	Malaria Vaccine Initiative
NMR	nuclear magnetic resonance
nt	nucleotide
OD	optical density
PABA	p-amino benzoic acid
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PPPK	hydroxymethyl pyrophosphokinase
PYR	pyrimethamine
RBS	ribosome binding site
SDS	sodium dodecylsulfate
SDX	sulfadoxine
SERCA	sarcoplasmic-endoplasmic reticulum Ca^{2+} ATPase
SHMT	serine hydroxymethyltransferase

SNP	single nucleotide polymorphism
TEMED	N,N,N'N'-tetramethylenediamine
Tris	tris (hydroxymethyl) aminomethane
tRNA	transcription ribonucleic acid
TS	thymidylate synthetase
U	enzyme units
UTR	untranslated region
X-gal	5-bromo-4-chloro-3-indolyl--β-D-galactopyranoside

List of tables

Chapter 1

Table 1.1: Problems associated with antimalarial drugs. 7

Table 1.2: Primary mutations responsible for antifolate resistance. 12

Chapter 2

Table 2.1: Forward (f) and reverse (r) oligonucleotides and forward (F) and reverse (R) primers used for the assembly of different gene fragments. 21

Table 2.2: Forward (F) and reverse (R) primers used for the generation of the quarter fragments. 34

Table2.3: Site-directed mutagenesis primers designed for error corrections in quarter 4. 35

Table 2.4: Forward (2f/3f) and reverse (2r/3r) oligonucleotides used for the resynthesis of the quarter 2 and 3 internal fragments. 38

Table 2.5: The theoretical and actual number of cycles needed for complete template generation based on the number of oligonucleotides involved. 43

Chapter 3

Table 3.1: Vector systems used for the recombinant expression of synthetic *P. falciparum dhfs-fpgs*. 56

Table 3.2: Primers used for cloning of *P. falciparum dhfs-fpgs* into the pASK-IBA3 vector for C-terminal Strep-tagged expression. 57

Table 3.3: *E. coli* strains used as hosts for protein expression. 57

List of figures

Chapter 1

Figure 1.1: The life cycle of the human malaria parasite *P. falciparum.* 2

Figure 1.2: The three main links in cycle of malaria transmission. 3

Figure 1.3: The subcellular locations of past, present and future drug targets. 8

Figure 1.4: Folate metabolism in the human malaria parasite *P. falciparum.* 11

Chapter 2

Figure 2.1: Schematic representation of the genetic organisation of *P. falciparum dhfs-fpgs*. 18

Figure 2.2: Overlap PCR used during gene synthesis. 20

Figure 2.3: Alignment of the synthetic and native (*Pf dhfs-fpgs*) *dhfs-fpgs* genes. 24

Figure 2.4: Overlap-extension PCR. 26

Figure 2.5: Subdivision of *dhfs-fpgs* into quarter segments with unique restriction enzyme sites in the overlaps. 34

Figure 2.6: Cassette replacement of a mutated gene segment with a newly synthesised internal area of the quarter fragment. 37

Figure 2.7: Graphs indicating the nucleotide composition of the synthetic *P. falciparum dhfs-fpgs* (A) gene compared to native *P. falciparum dhfs-fpgs* (B). 40

Figure 2.8: Taguchi optimisation of the Mg^{2+}, dNTP and *Pfu* DNA polymerase concentrations for the assembly of six oligonucleotides. 41

Figure 2.9: Refinement of the PCR parameters for the assembly PCR step. 42

Figure 2.10: PCR products obtained for fragments assembled from 6 oligonucleotides (A) or 10 oligonucleotides (B). 43

Figure 2.11: The full-length synthetic *dhfs-fpgs* gene obtained from overlap extension PCR of the two 750 bp half fragments. 44

Figure 2.12: The direct relationship between the error rate and the proximity to the 5' end of each oligonucleotide. 45

Figure 2.13: Construction of the correct quarter 2 by means of cassette mutagenesis. 45

Figure 2.14: Corrected quarter fragments. 46

Figure 2.15: Isolation of the 1100bp fragment consisting of the 1st three fragments by means of restriction digestion. 47

Figure 2.16: Full-length *dhfs-fpgs* constructs obtained after overlap-extension of the 1100bp fragment and alternative 500bp fragments. 47

Chapter 3

Figure 3.1: Expression of (A) pET15b-*dhfs-fpgs* (N-terminal His_6 tag) and (B) pET22b-*dhfs-fpgs* (C terminal His_6 tag) in various cell lines. 62

Figure 3.2: Solubility of C-terminal His-tagged DHFS-FPGS in A: BL21 Star (DE3), B: BL21 Gold (DE3) pLysS and C: BL21 (DE3) pLysS cell lines. 63

Figure 3.3: Solubility of tagless DHFS-FPGS in A: BL21 Star (DE3), B: BL21 Gold (DE3) pLysS and C: BL21 (DE3) pLysS cell lines. 63

Figure 3.4: Expression of C-terminal Strep-tagged *dhfs-fpgs* in A: BL21 Star (DE3), B: BL21 Gold (DE3) pLysS and C: BL21 (DE3) pLysS cell lines. 64

Figure 3.5: Complementation of DHFS-FPGS deficient *E. coli* (SF4) by different synthetic *P. falciparum dhfs-fpgs* constructs. 64

Figure 3.6: Growth of SF4 *E.coli* containing different constructs in liquid media. 65

Figure 3.7 Growth curves over 24 hours for the tagless *dhfs-fpgs* construct in Sf4 cells (tl) versus the negative control Sf4 cells without any construct (Sf4). 66

Figure 3.8: Affinity purification of resolubilised C-terminal His_6-tagged DHFS-FPGS obtained from BL21 Star (DE3) cells. 66

Figure 3.9: C-terminal His_6-tagged DHFS-FPGS affinity purified from the soluble fraction expressed by BL21 (DE3) pLysS cells. 67

Figure 3.10: Retention times of low molecular mass protein standards. 68

Figure 3.11: Size exclusion analysis of the expression of C-terminal His_6-tagged DHFS-FPGS from BL21 (DE3) pLysS cells. 68

Figure 3.12: Size exclusion HPLC profiles obtained for affinity purified C-terminal His_6-tagged DHFS-FPGS from BL21 (DE3) pLysS cells. 69

Figure 3.13: Silver stained SDS PAGE of fractions obtained after size exclusion HPLC of C-terminal His_6-tagged DHFS-FPGS. 69

Chapter 4

Figure 4.1: DHFS and FPGS enzyme reactions. 74

Figure 4.2: Schematic representation of *L. casei* FPGS. 75

Figure 4.3: Phylogenetic analysis of *P. falciparum* DHFS-FPGS based on its alignment with homologous proteins. 79

Figure 4.4: Alignment of *P. falciparum* DHFS-FPGS with homologous proteins. 84

Figure 4.5: Superimposed ribbon backbones of the preliminary homology model of *P. falciparum* DHFS-FPGS on the *L. casei* crystal structure. 85

Figure 4.6: Ramachandran plot of the *P. falciparum* DHFS-FPGS homology model. 85

Figure 4.7: *Plasmodium* DHFS-FPGS vs human FPGS alignment. 87

Figure 4.8: Secondary structure prediction of *P. falciparum* DHFS-FPGS by GOR4. 88

Figure 4.9: Alignment of independent secondary structure predictions. 89

Figure 4.10: Hydrophobicity profile for the primary amino acid sequence using Kyte and Doolittle parameters. 91